

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2002, 22:40:30 ; Search time 304.605 Seconds
(without alignments)
11355.910 Million cell updates/sec

Title: US-09-963-333-6

Perfect score: 1536

Sequence: 1 ggggggggggggaccacttg.....ataataaagaagtgtctgc 1536

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched:2185239 seqs, 1125999159 residues

Word size: 15

Total number of hits satisfying chosen parameters: 4914

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

N_Geneseq_101002.*

1: /SID82/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID82/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID82/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID82/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID82/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID82/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID82/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID82/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID82/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID82/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID82/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID82/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID82/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID82/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID82/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID82/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID82/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID82/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID82/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID82/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID82/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID82/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID82/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID82/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1394	90.8	1536	ABK43335	Human Thymidylate
2	1346	87.6	3298	AAS94945	Human DNA sequence
3	536	34.9	1539	AAS84960	DNA encoding novel
4	532	34.6	535	ABL38559	Human colon tumour
5	487	31.7	18596	AAF31109	Thymidylate synthase
6	487	31.7	18596	AAC91215	Human Thymidylate
7	487	31.7	18596	ABN95092	Gene #1590 used to
8	487	31.7	18596	ABK43334	Human Thymidylate
9	487	31.7	18596	ABL62854	Breast cancer rela

10	487	31.7	18596	24	ABL63078	Breast cancer rela
11	487	31.7	18596	24	ABL67927	Ovary cancer relat
12	487	31.7	45716	24	ABA93401	Human rts-alpha ge
13	487	31.7	45989	24	ABA93402	Human rts-beta gen
14	379	24.7	566	20	AA24270	Human thymidylate
15	178	11.6	651	23	AAS84956	DNA encoding novel
16	168	10.9	1817	24	ABA93399	Human rts-alpha en
17	168	10.9	1857	22	AA158987	Human polynucleoti
18	150	9.8	196	22	AAS04589	Gene expression pr
19	142	9.2	346	21	AA78156	cDNA encoding huma
20	142	9.2	346	22	AA128894	Colon tumour relat
21	107	7.0	1131	24	ABK43330	Human GNEHK genomi
22	107	7.0	1161	24	ABK43288	Human CDNA encodin
23	79	5.1	118	19	AA10927	Human biallelic po
24	77	5.0	1692	22	ABA09660	Human bone marrow
25	77	5.0	3422	23	AAS84959	DNA encoding novel
26	60	3.9	60	24	ABN32855	Human spliced tran
27	59	3.8	196	22	AAS04572	Gene expression pr
28	40	2.6	2945	23	AAS84962	DNA encoding novel
29	30	2.0	30	20	AA24268	Thymidylate synthase
30	30	2.0	30	20	AA24269	Thymidylate synthase
31	30	2.0	30	20	AAV71964	Thymidylate synthase
32	30	2.0	30	20	AAV71965	Thymidylate synthase
33	30	2.0	36	20	AA24274	Thymidylate synthase
34	30	2.0	36	20	AA24271	Thymidylate synthase
35	30	2.0	36	20	AA24272	Thymidylate synthase
36	30	2.0	36	20	AA24273	Thymidylate synthase
37	30	2.0	230	23	ABV57305	Human prostate exp
38	29	1.9	29	22	AAH41174	Human thymidylate
39	27	1.8	764	24	ABQ39528	Oligonucleotide fo
40	27	1.8	764	24	ABQ39529	Oligonucleotide fo
41	27	1.8	1186	24	ABK40090	Human chemically p
42	26	1.7	26	19	AAK09473	Human biallelic po
43	26	1.7	26	19	AAV52596	Primer hts-3a, use
44	26	1.7	128139	24	AA164291	RRV genome nucleot
45	26	1.7	133719	21	AA64754	Macaca mulatta rha

ALIGNMENTS

RESULT 1

ABK43335 ID ABK43335 standard; cDNA; 1536 BP.

XX AC ABK43335;

XX 05-JUN-2002 (first entry)

XX DT Human Thymidylate synthase cDNA sequence.

XX DE HKNG1; ss; gene; chromosome 18p; bipolar affective disorder; BAD;
XX KW severe bipolar affective (mood) disorder; BP-I; schizophrenia; TS;
XX KW Hong Kong new gene 1; antimanic; antidepressant; neuroleptic;
XX KW Thymidylate synthase.

XX OS Homo sapiens.

XX PN WO200210366-A2.

XX PD 07-FEB-2002.

XX PF 02-AUG-2001; 2001WO-US24417.

XX PR 02-AUG-2000; 2000US-0631275.

XX PR 28-NOV-2000; 2000US-0722544.

XX PA (MILL-) MILLENNIUM PHARM INC.
(REGC) UNIV CALIFORNIA.

XX PI Chen H, Freimer NB, Novak T;

XX DR WPI; 2002-195962/25.

Db 52 CACCATCGATCATGATGATAGATGTGGTTGTGAACTTTATCGTGTGTTTATA 1
|||||

RESULT 4
AL547612 1020 bp mRNA linear EST 16-FEB-2001
LOCUS
DEFINITION AL547612 LTI_NFL006_PL2 Homo sapiens cDNA clone CSOD1008YH21 5
prime, mRNA sequence.
ACCESSION AL547612
VERSION AL547612.1 GI:12881832
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1020)
REFERENCE
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope -- Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 1020
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSOD1008YH21"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site: 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Peng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : filiang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 237 a 291 c 276 g 215 t 1 others
ORIGIN
Query Match 66.3%; Score 1018; DB 9; Length 1020;
Best Local Similarity 99.8%; Pred. No. 2.4e-254;
Matches 1018; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 35 CGCGCGCGCACTTGGCGCTCCCTCCGCGCGCGCGCACTTCGCTGCCTCCGTCGCCCG 94
DB 1 CGCGCGCGCACTTGGCGCTCCCTCCGCGCGCGCGCACTTCGCTGCCTCCGTCGCCCG 60
QY 95 CGCGCGCGCGCATGCTGTGGCGCGCTCGAGAGTGGCGCGCGCGCTTGGCCCCCGCG 154
DB 61 CGCGCGCGCGCATGCTGTGGCGCGCTCGAGTGGCGCGCGCGCTTGGCCCCCGCG 120
QY 155 CACAGACGGGACGCGGACGCGCTTCGCGCGCACGGGAGCTGCAGTACTGGGGCAGA 214
DB 121 CACAGAGCGGGACGCGGACGCGCGCTTCGCGCGCACGGGAGCTGCAGTACTGGGGCAGA 180
QY 215 TCCACACATCTCCGCTGGCGGCTCAGGAAGGACGACCGCAGGCGCGCGCGCGCGCGCTGT 274
DB 181 TCCACACATCTCCGCTGGCGGCTCAGGAAGGACGACCGCAGGCGCGCGCGCGCGCGCTGT 240
QY 275 CGGTATTCGGCATCGAGGCGCGCTACACCTTGAGAGATGAATTCCTCTGCTGCACACCA 334
DB 241 CGGTATTCGGCATCGAGGCGCGCTACACCTTGAGAGATGAATTCCTCTGCTGCACACCA 300
QY 335 AACGCTGTCTTCGAAGGGTGTTTGGAGGAGTGTCTGTGGTTATCAAGGGATCCACAA 394
DB 301 AACGCTGTCTTCGAAGGGTGTTTGGAGGAGTGTCTGTGGTTATCAAGGGATCCACAA 360
QY 395 ATGCTAAAGAGCTGTCTTCCAAAGGGAGTGAAAATCTGGGATCCGAATGGATCCCGAGACT 454

JOURNAL COMMENT

Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

Location/Qualifiers
1. .1089
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODI061YF09"
/clone_lib="LRI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 298 a 253 c 217 g 318 t 3 others

ORIGIN

Query Match 68.2%; Score 1048.2; DB 9; Length 1089;
Best local Similarity 98.9%; Pred. No. 3.2e-262;
Matches 1072; Conservative 2; Mismatches 8; Indels 2; Gaps 2;

QY 432 GGATGCCAATGATCCGAGACTTTTGGACAGCTTGGACAGCTGGGATTCACACAGAGAAGG 491
DB 1083 GGATGCCAATGATCCGAGACTTTTGGACAGCTTGGACAGCTGGGATTCACACAGAGA 1026
QY 492 GGACTTGGCCCGAGTTATAGCTTCCAGTGGAGGCATTTTGGGGCAGAAATACAGAGATAT 551
DB 1025 GGACTTGGCCCGAGTTATAGCTTCCAGTGGAGGCATTTTGGGGCAGAAATACAGAGATAT 966
QY 552 GGAATCAGATTATTCAGGACAGGAGTTGACCACTGCAAGAGTGATTCACACATCAA 611
DB 965 GGAATCAGATTATTCAGGACAGGAGTTGACCACTGCAAGAGTGATTCACACATCAA 906
QY 612 AACCAACCTTGACAGAGAAATCATCATGTGGCTTGGAAATCCAGAGATCTTCTCT 671
DB 905 AACCAACCTTGACAGAGAAATCATCATGTGGCTTGGAAATCCAGAGATCTTCTCT 846
QY 672 GATGGCGCTGCTCCATGCGCATGCGCTCTGCCAGTTCTATGTGTGTCACACAGCTGCTGTC 731
DB 845 GATGGCGCTGCTCCATGCGCATGCGCTCTGCCAGTTCTATGTGTGTCACACAGCTGCTGTC 786
QY 732 CTGCGAGCTGTACAGAGATCGGAGACATGGGCGCTCGGTGTGCTTTCAACATCGCCAG 791
DB 785 CTGCGAGCTGTACAGAGATCGGAGACATGGGCGCTCGGTGTGCTTTCAACATCGCCAG 726
QY 792 CTAGCGCTGCTCAGTACATGATTTGGGACATCAGGGCGCTGAAGCAGCTGATCTTAT 851
DB 725 STAGCGCTGCTCAGTACATGATTTGGGACATCAGGGCGCTGAAGCAGCTGATCTTAT 666
QY 852 ACACATTTGGGAGATGCACATATTTACCTGATTCATCATCGGCGCTGCAAGATTCAGCT 911
DB 665 ACACATTTGGGAGATGCACATATTTACCTGATTCATCATCGGCGCTGCAAGATTCAGCT 606
QY 912 TCAGCGAGAACCCAGACCTTTCCCAAGCTTCAGATTCTTCGAAAGTTGAGAAATTTGA 971
DB 605 TCAGCGAGAACCCAGACCTTTCCCAAGCTTCAGATTCTTCGAAAGTTGAGAAATTTGA 546
QY 972 TGACTTCAAGCTTGAAGACTTTCAGATTGAAGGTGTAATCCGATCCCACTTAAAT 1031
DB 545 TGACTTCAAGCTTGAAGACTTTCAGATTGAAGGTGTAATCCGATCCCACTTAAAT 486
QY 1032 GGAATAGCTGTTTAGGGTGCTTTCAAGGAGCTNGAAGATATTGTGCTTTAGGGG 1091
DB 485 GGAATAGCTGTTTAGGGTGCTTTCAAGGAGCTTTGAAGATATTGTGCTTTAGGGG 426

QY 1092 TTGGCTGGATGCGAGGTAAAGTTCTTTTGTCTTAAAGAAAGAAAGAACTAGTCAA 1151
DB 425 TTGGCTGGATGCGAGGTAAAGTTCTTTTGTCTTAAAGAAAGAAAGAACTAGTCAA 366
QY 1152 AAATCTGTCCTGACCTATCAGTTATTAATTTTAAAGCATCTTCCCACTGCGCAAAATGTA 1211
DB 365 AAATCTGTCCTGACCTATCAGTTATTAATTTTAAAGCATCTTCCCACTGCGCAAAATGTA 306
QY 1212 CTGTCCGAGTTCTTTCCATTAATAAAGGCTTTGAGTTAACTCACTGAGGGTATCTGACAA 1271
DB 305 CTGTCCGAGTTCTTTCCATTAATAAAGGCTTTGAGTTAACTCACTGAGGGTATCTGACAA 246
QY 1272 TCCTCAGGTATGACAAAGTGGAGAGATGAATGATATGTGCTCTTAGCAAAAACATGT 1331
DB 245 TCCTCAGGTATGACAAAGTGGAGAGATGAATGATATGTGCTCTTAGCAAAAACATGT 186
QY 1332 ATGTGCAATTTCAATCCACGACTTATTAATAAAGGCTTTGAGTTAACTCACTGAGGGTATTTT 1391
DB 185 ATGTGCAATTTCAATCCACGACTTATTAATAAAGGCTTTGAGTTAACTCACTGAGGGTATTTT 126
QY 1392 GGAATATTTTGAATATTTTAAAGATTTTCAAGCTATTTCCCTCAAACTCGAGGAGCT 1451
DB 125 GGAATATTTTGAATATTTTAAAGATTTTCAAGCTATTTCCCTCAAACTCGAGGAGCT 66
QY 1452 GAGTAACACCATCGATCATGATGTAGAGTGTGTTATCAAACTTTANAGTTGTTTATATG 1511
DB 65 GAGTAACACCATCGATCATGATGTAGAGTGTGTTATCAAACTTTANAGTTGTTTATATG 6
QY 1512 TTGC 1515
DB 5 TTGC 2

RESULT 3

AL572496/c

LOCUS

DEFINITION AL572496 LRI_NFL006_PL2 Homo sapiens cDNA clone CS0D1008YH21 3
prime, mRNA sequence.

ACCESSION AL572496

VERSION AL572496.1 GI:12930822

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1077)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. .1077

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CSODI008YH21"

/clone_lib="LRI_NFL006_PL2"

/tissue_type="placenta"

/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact : Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 294 a 251 c 214 g 309 t 9 others

ORIGIN

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

M nucleic - nucleic search, using sw model

un on: November 28, 2002, 19:35:09 ; Search time 2460.31 Seconds
(without alignments)
12130.226 Million cell updates/sec

itle: US-09-963-333-7

effect score: 1187

equences: 1 gatcgccactgactcca.....gagggagcgcgctgggga 1187

coring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

earched: 24791104 seqs, 12571243825 residues

otal number of hits satisfying chosen parameters: 49582208

inimum DB seq length: 0

aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase : Pending Patents NA Main.*

1: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*

44: /cgn2_6/ptodata/2/pna/US000 COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US001 COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US002 COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US003 COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US004 COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US005 COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US006 COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US007 COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US008 COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US009 COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US010 COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US011 COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US012 COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US013 COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US014 COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US015 COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US016 COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US017 COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US018 COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US019 COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US020 COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US021 COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US022 COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US023 COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US024 COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US025 COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US026 COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US027 COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US028 COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US029 COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US030 COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US031 COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US032 COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US033 COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US034 COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US035 COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US036 COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US037 COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US038 COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US039 COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US040 COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US041 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1178	99.2	1187	25	US-09-658-659-7	Sequence 7, Appli
2	1178	99.2	1187	36	US-09-963-665-7	Sequence 7, Appli
3	1178	99.2	1187	36	US-09-963-677-7	Sequence 7, Appli
4	1178	99.2	1187	36	US-09-963-333-7	Sequence 7, Appli
5	1167	98.3	18596	22	US-09-577-266-11	Sequence 11, Appli
6	1167	98.3	18596	24	US-09-631-275-140	Sequence 140, App
7	1167	98.3	18596	33	US-09-880-107-1590	Sequence 1590, App
8	1167	98.3	18596	36	US-09-954-531-124	Sequence 124, App
9	1167	98.3	18596	36	US-09-954-531-348	Sequence 348, App
10	1167	98.3	18596	36	US-09-967-768A-119	Sequence 119, App
11	1167	98.3	18597	25	US-09-658-659-8	Sequence 8, Appli
12	1167	98.3	18597	36	US-09-962-665-8	Sequence 8, Appli
13	1167	98.3	18597	36	US-09-963-333-8	Sequence 8, Appli
14	1167	98.3	18597	36	US-09-963-333-8	Sequence 9676, Ap
15	335	28.2	1829	25	US-09-652-109-9676	Sequence 6866, Ap
16	335	28.2	1829	25	US-09-652-121-6686	Sequence 8870, Ap
17	335	28.2	1829	25	US-09-652-816-8870	Sequence 6995, Ap
18	335	28.2	1829	27	US-09-698-013-6995	Sequence 2244, Ap
19	335	28.2	1829	28	US-09-716-953-2244	Sequence 4489, Ap
20	335	28.2	1829	28	US-09-717-350-4489	Sequence 5305, Ap
21	335	28.2	1829	29	US-09-721-588-5305	Sequence 5305, Ap

22	335	28.2	1829	29	US-09-726-174-5178
23	335	28.2	1829	29	US-09-726-806-5390
24	335	28.2	1829	29	US-09-726-807-3895
25	299.4	25.2	1536	1	PCT-US02-18947-556
26	299.4	25.2	1536	1	PCT-US99-12080-1
27	299.4	25.2	1536	18	US-09-442-384A-792
28	299.4	25.2	1536	18	US-09-442-384B-792
29	299.4	25.2	1536	24	US-09-631-275-141
30	299.4	25.2	1536	25	US-09-658-659-6
31	299.4	25.2	1536	33	US-09-877-178-11
32	299.4	25.2	1536	36	US-09-962-663-6
33	299.4	25.2	1536	36	US-09-962-677-6
34	299.4	25.2	1536	36	US-09-963-333-6
35	299.4	25.2	1536	41	US-10-172-118-556
36	299.4	25.2	1536	85	US-60-412-049-350
37	299.4	25.2	1540	1	PCT-US01-14827-3952
38	299.4	25.2	1540	22	US-09-577-408-4887
39	299.4	25.2	2023	22	US-09-572-411-5049
40	288.8	24.3	3298	63	US-60-195-106-200
41	288.8	24.3	3298	65	US-60-217-674-28
42	288.8	24.3	3298	71	US-60-278-258-674
43	273.8	23.1	464	17	US-09-399-932-2144
44	273.8	23.1	1633	61	US-60-172-373-69
45	264	22.2	437	29	US-09-726-806-2923

ALIGNMENTS

RESULT 1
US-09-658-659-7
; Sequence 7, Application US/09658659
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: GENE SEQUENCE VARIANCE IN GENES RELATED
; TITLE OF INVENTION: TO FOLATE METABOLISM HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015001
; CURRENT APPLICATION NUMBER: US/09/658,659
; CURRENT FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 276, 321, 534, 656
; OTHER INFORMATION: n = c or t
; NAME/KEY: misc feature
; LOCATION: 452, 640
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc feature
; LOCATION: 492-625
; OTHER INFORMATION: n = c or a
; NAME/KEY: misc feature
; LOCATION: 458
; OTHER INFORMATION: nucleotide in position 458 is c, or absent
US-09-658-659-7

Query Match 99.2%; Score 1178; DB 25; Length 1187;
Best Local Similarity 100.0%; Pred. No. 5.2e-117;
Matches 1187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GATCGCGCACTGCATCCAGCGCTGGGTGAGAGAGCGAGACTCTGTCTCAAAAAA	60
Db	1	GATCGCGCACTGCATCCAGCGCTGGGTGAGAGAGCGAGACTCTGTCTCAAAAAA	60
Qy	61	AAAAAGACGGCAGCGCTCAAAACAAAAACCTTCGAAAAAGCCCTGGCGGTCTTTTTTTT	120
Db	61	AAAAAGACGGCAGCGCTCAAAACAAAAACCTTCGAAAAAGCCCTGGCGGTCTTTTTTTT	120
Qy	121	TTTTTTTTTTTTTTTTTTTTTTGGGACAGTCTTGTCTGTGCGCCAGGCTGGAGTACAATGT	180
Db	121	TTTTTTTTTTTTTTTTTTTTTTGGGACAGTCTTGTCTGTGCGCCAGGCTGGAGTACAATGT	180
Qy	181	CGGATCTTGGCTCACTGGAACCTCTGCTCCAGGTTCAAGCAATTTCTTGCTCAGCC	240
Db	181	CGGATCTTGGCTCACTGGAACCTCTGCTCCAGGTTCAAGCAATTTCTTGCTCAGCC	240
Qy	241	TCCCAAGTAGACACACGCCAGCTAAATTTTGTANTTTAGTAGAGACGGGGTTTCA	300
Db	241	TCCCAAGTAGACACACGCCAGCTAAATTTTGTANTTTAGTAGAGACGGGGTTTCA	300
Qy	301	CATGTTGTCAGGCTGCTGTAACCTCTGATCCAGGTCACCCGCTCGGCCCCC	360
Db	301	CATGTTGTCAGGCTGCTGTAACCTCTGATCCAGGTCACCCGCTCGGCCCCC	360
Qy	361	CAAAAGTACTAGGATTAAGCGGTGAGCCACCGCTCCAGCGCCCTGGCGGTCTTTAATCA	420
Db	361	CAAAAGTACTAGGATTAAGCGGTGAGCCACCGCTCCAGCGCCCTGGCGGTCTTTAATCA	420
Qy	421	AGTAGAAAAGCTGCATTATACCACTTCTTCNGTTTCNNTTCAGTGAGAAACGAAGAAATGG	480
Db	421	AGTAGAAAAGCTGCATTATACCACTTCTTCNGTTTCNNTTCAGTGAGAAAGAAATGG	480
Qy	481	AAATGCAATNCNTTATTTAGTTAGAGAAACAGATCTCAACACAGCTTTTGTNGACAG	540
Db	481	AAATGCAATNCNTTATTTAGTTAGAGAAACAGATCTCAACACAGCTTTTGTNGACAG	540
Qy	541	ACCGAGGAAAACGTGGGAACCTGCTGCTGCTTAGAGAGGCGCGCTCGACACAGCG	600
Db	541	ACCGAGGAAAACGTGGGAACCTGCTGCTGCTTAGAGAGGCGCGCTCGACACAGCG	600
Qy	601	TTCCCAAGGCGCAGTCTTCCNGCCACCGCACTGCTCCAGGTTCCCGGTTCCT	660
Db	601	TTCCCAAGGCGCAGTCTTCCNGCCACCGCACTGCTCCAGGTTCCCGGTTCCT	660
Qy	661	AAGACTCTCAGCTGTGGCCCTTGCGCTCCGTTCTGTGCCACACCCCGTGGCTTCCTCGCTTC	720
Db	661	AAGACTCTCAGCTGTGGCCCTTGCGCTCCGTTCTGTGCCACACCCCGTGGCTTCCTCGCTTC	720
Qy	721	CCCTCGGCGCACGCTCTTAGAGCGGGGGCGCGCGACCCCGCGCAGAGAGAGGGCG	780
Db	721	CCCTCGGCGCACGCTCTTAGAGCGGGGGCGCGCGACCCCGCGCAGAGAGAGGGCG	780
Qy	781	GAGCGGGGACGGCGCGGAAAAGGGCGCGGAGAGGGGTCTGCGACCCGCGCACTTGG	840
Db	781	GAGCGGGGACGGCGCGGAAAAGGGCGCGGAGAGGGGTCTGCGACCCGCGCACTTGG	840
Qy	841	CTGCTCGCTCCGTCGCGCGCACTTGGCTCCGTCCTCGCGCGCGCCACTTTCGCTG	900
Db	841	CTGCTCGCTCCGTCGCGCGCACTTGGCTCCGTCCTCGCGCGCGCCACTTTCGCTG	900
Qy	901	CTGCTCGCTCCGTCGCGCGCGCATGCTGTGGCGCGGTGCGAGCTGCGCGCGCGCC	960
Db	901	CTGCTCGCTCCGTCGCGCGCGCATGCTGTGGCGCGGTGCGAGCTGCGCGCGCGCC	960
Qy	961	TTGCCCCCGCGCACAGGAGCGCGGACCGCGCGCTGCTCCCGCCACAGGGAGCTGCGAG	1020
Db	961	TTGCCCCCGCGCACAGGAGCGCGGACCGCGCGCTGCTCCCGCCACAGGGAGCTGCGAG	1020
Qy	1021	TACTTGGGCGACATCCAAACATCTCTCGCTGCGCGCTCAGAGAGAGAGACCGGCG	1080
Db	1021	TACTTGGGCGACATCCAAACATCTCTCGCTGCGCGCTCAGAGAGAGAGACCGGCG	1080
Qy	1081	ACCGCGACCCCTGTCGTTATTCGGCATTCAGGCGCGCTACGCTGAGAGGTGAGCGCG	1140

OTHER INFORMATION: n = c or t
NAME/KEY: misc_feature
LOCATION: 452, 640
OTHER INFORMATION: n = a or g
NAME/KEY: misc_feature
LOCATION: 492_625
OTHER INFORMATION: n = c or a
NAME/KEY: misc_feature
LOCATION: 458
OTHER INFORMATION: nucleotide in position 458 is c, or absent
US-09-963-333-7

Query Match 99.2%; Score 1178; DB 36; Length 1187;
Best Local Similarity 100.0%; Pred. No. 5.2e-117; Indels 0; Gaps 0;
Matches 1187; Conservative 0; Mismatches 0;

1 GATCGCGCCACTGCACTCCAGCGCTGGGTGAGAGCGAGAGCTCTGCTCAAAAAA 60
1 GATCGCGCCACTGCACTCCAGCGCTGGGTGAGAGCGAGAGCTCTGCTCAAAAAA 60
61 AAAAAAGCGCGCAGGCGCTCAAAACAAAAACCTCGAAAAAGCCCTGGCGGTCTTTT 120
61 AAAAAAGCGCGCAGGCGCTCAAAACAAAAACCTCGAAAAAGCCCTGGCGGTCTTTT 120
121 TTTTCTTTTCTTTTCTTTTGGGCAAGCTCTCTCTGCGCCAGGCTGGAGTACAATGT 180
121 TTTTCTTTTCTTTTCTTTTGGGCAAGCTCTCTCTGCGCCAGGCTGGAGTACAATGT 180
181 CGGATCTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTTCCTCAGCC 240
181 CGGATCTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTTCCTCAGCC 240
241 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTTGTAGAGAGCGGGTTTCA 300
241 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTTGTAGAGAGCGGGTTTCA 300
301 CATGTTCTCCAGGCTGGTCTGAGTCTCTGACCTCAGGTGATCCACCCGCTCGGCC 360
301 CATGTTCTCCAGGCTGGTCTGAGTCTCTGACCTCAGGTGATCCACCCGCTCGGCC 360
361 CAAGTACTAGGATTAACAGCGTGAAGCCACCGCTGACCGCCCTCGCGGTCTTAA 420
361 CAAGTACTAGGATTAACAGCGTGAAGCCACCGCTGACCGCCCTCGCGGTCTTAA 420
421 AGTAGAAAAGTGCATTAATACCACTTGTCTGTTGTTGTTGTTGTTGTTGTTGTT 480
421 AGTAGAAAAGTGCATTAATACCACTTGTCTGTTGTTGTTGTTGTTGTTGTTGTT 480
481 AAATCAAAATCNCCTTATGTTAGTAAACAGATCTCAACACAGCTTTTGTGACAG 540
481 AAATCAAAATCNCCTTATGTTAGTAAACAGATCTCAACACAGCTTTTGTGACAG 540
541 ACCGAGAAAACCTGGGAATCTGCTGCTGCTTTAGAGAGGCGGCTCGACAGAGG 600
541 ACCGAGAAAACCTGGGAATCTGCTGCTGCTTTAGAGAGGCGGCTCGACAGAGG 600
601 TTCCCAAGAGGCGAGTCTCTCCCGCCAGCCAGCTGCTCCAGGTTCCCGGTTTCT 660
601 TTCCCAAGAGGCGAGTCTCTCCCGCCAGCCAGCTGCTCCAGGTTCCCGGTTTCT 660
661 AAGACTCTCAGCTGTGGGCTTGGGCTCCGTTCTGTGCGCCACACCCGCTGGCTTTC 720
661 AAGACTCTCAGCTGTGGGCTTGGGCTCCGTTCTGTGCGCCACACCCGCTGGCTTTC 720
721 CCCTGTGGCGCAGCTCTCTAGAGCGGGGCGCGCGACCCCGCGCAGAGAGAGGCG 780
721 CCCTGTGGCGCAGCTCTCTAGAGCGGGGCGCGCGACCCCGCGCAGAGAGAGGCG 780
781 GAGCGGGGAGCGCGCGGGGAAAAGCGCGCGAGAGGGGTCTGCGCACCGGCACTTGG 840
781 GAGCGGGGAGCGCGCGGGGAAAAGCGCGCGAGAGGGGTCTGCGCACCGGCACTTGG 840
841 CTTGCTCTCCGTCGCCGCGGCCACTTGGGCTGCTCCGTCGCCGCGGCCACTTGGCTG 900

Db 841 CTTGCTCTCCGTCGCCGCGGCCACTTGGCTGCTCCGTCGCCGCGGCCACTTGGCTG 900
Qy 901 CTTGCTCTCCGTCGCCGCGGCCACTTGGCTGCTCCGTCGCCGCGGCCACTTGGCTG 960
Db 901 CTTGCTCTCCGTCGCCGCGGCCACTTGGCTGCTCCGTCGCCGCGGCCACTTGGCTG 960
Qy 961 TTGCTCTCCGTCGCCGCGGCCACTTGGCTGCTCCGTCGCCGCGGCCACTTGGCTG 1020
Db 961 TTGCTCTCCGTCGCCGCGGCCACTTGGCTGCTCCGTCGCCGCGGCCACTTGGCTG 1020
Qy 1021 TACCTGGGGCAGATCAACACATCTCTGCTGCGGCTCAGGAAGGACGACGACGCGGC 1080
Db 1021 TACCTGGGGCAGATCAACACATCTCTGCTGCGGCTCAGGAAGGACGACGACGCGGC 1080
Qy 1081 ACCGCGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 1081 ACCGCGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Qy 1141 GGGCTCTGCGGAGCGGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1187
Db 1141 GGGCTCTGCGGAGCGGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1187

RESULT 5
US-09-577-266-11
; Sequence 11, Application US/09577266
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Steinhilber, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; TITLE OF INVENTION: DEVELOPMENTAL DISORDERS
; FILE REFERENCE: 601-1-057N
; CURRENT APPLICATION NUMBER: US/09577266
; CURRENT FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/136,198
; PRIOR FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 11
; LENGTH: 18596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-577-266-11

Query Match 98.3%; Score 1167; DB 22; Length 18596;
Best Local Similarity 99.2%; Pred. No. 4.5e-116;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Qy 1 GATCGCGCCACTGCACTCCAGCGCTGGGTGAGAGCGAGAGCTCTGCTCAAAAAA 60
Db 78 GATCGCGCCACTGCACTCCAGCGCTGGGTGAGAGCGAGAGCTCTGCTCAAAAAA 137
Qy 61 AAAAAAGCGCGCAGGCTCAAAACAAAAACCTCGAAAAAGCCCTGGCGGTCTTTT 120
Db 138 AAAAAAGCGCGCAGGCTCAAAACAAAAACCTCGAAAAAGCCCTGGCGGTCTTTT 197
Qy 121 TTTTCTTTTCTTTTCTTTTGGGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 198 TTTTCTTTTCTTTTCTTTTGGGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 257
Qy 181 CGGATCTGGCTCACTGCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 258 CGGATCTGGCTCACTGCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 317
Qy 241 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTTGTAGAGAGCGGGTTTCA 300
Db 318 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTTGTAGAGAGCGGGTTTCA 377
Qy 301 CATGTTCTCCAGGCTGGTCTGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 378 CATGTTCTCCAGGCTGGTCTGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 437

QY	721	CCCTGGCGCACGCTCTCTAGAGCGGGGGCGCGCGACCCCGCGAGCAGGAAGAGCGG	780
DB	797	CCCTGGCGCACGCTCTCTAGAGCGGGGGCGCGCGACCCCGCGAGCAGGAAGAGCGG	856
QY	781	GAGCGGGGACGCGCGCGGGAAGCGCGGGAAGGGGFTCTGTGCCACCGCGCCACTTGG	840
DB	857	GAGCGGGGACGCGCGCGGGAAGCGCGGGAAGGGGFTCTGTGCCACCGCGCCACTTGG	916
QY	841	CCTGCCTCCGTCCTCCGCGCGGCCACATTGCGCTTCCGTCGCGCGCGCCACTTGCCTG	900
DB	917	CCTGCCTCCGTCCTCCGCGCGGCCACTTGCCTTCCGTCGCGCGCGCCACTTGCCTG	976
QY	901	CCTCGCTCCCGCGCGCGCGCGCATGCTGTGCGCGGCTCGAGAGTCCGCGCGCGGCC	960
DB	977	CCTCGCTCCCGCGCGCGCGCGCATGCTGTGCGCGGCTCGAGAGTCCGCGCGCGGCC	1036
QY	961	TTGCCCCCGCGCACAGAGCGGGAGCGCGCGCGGCTCCGCGCACGCGGGAGCTGCAG	1020
DB	1037	TTGCCCCCGCGCACAGAGCGGGAGCGCGCGCGGCTCCGCGCACGCGGGAGCTGCAG	1096
QY	1021	TACCTGGGGCAGATCAACACATCTCTCGCTGCGCGCTCAGGAAGCACACCGCACGCGC	1080
DB	1097	TACCTGGGGCAGATCAACACATCTCTCGCTGCGCGCTCAGGAAGCACACCGCACGCGC	1156
QY	1081	ACGGCACCTGTGCTGTTTGGCATGACAGCGCGCTACAGCTTGAGAGTGAAGCGCGG	1140
DB	1157	ACGGCACCTGTGCTGTTTGGCATGACAGCGCGCTACAGCTTGAGAGTGAAGCGCGG	1216
QY	1141	GCGCCCTCGGACGCGGTGCGGGAGAGGAGGCGCGCGCTGGGA	1187
DB	1217	GCGCCCTCGGACGCGGTGCGGGAGAGGAGGCGCGCGCTGGGA	1263

RESULT 11
 US-09-658-659-8
 ; Sequence 8, Application US/09658659
 ; GENERAL INFORMATION:
 ; APPLICANT: Scanton, Jr., Vincent P.
 ; TITLE OF INVENTION: GENE SEQUENCE VARIANCE IN GENES RELATED
 ; TITLE OF INVENTION: TO FOLATE METABOLISM HAVING UTILITY IN DETERMINING THE
 ; TITLE OF INVENTION: TREATMENT OF DISEASE
 ; FILE REFERENCE: 11926-015001
 ; CURRENT APPLICATION NUMBER: US/09/658,659
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 09/596,033
 ; PRIOR FILING DATE: 2000-06-15
 ; PRIOR APPLICATION NUMBER: 09/357,743
 ; PRIOR FILING DATE: 1999-07-20
 ; PRIOR APPLICATION NUMBER: 09/357,024
 ; PRIOR FILING DATE: 1999-07-19
 ; PRIOR APPLICATION NUMBER: 60/093,484
 ; PRIOR FILING DATE: 1998-07-20
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 18597
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 701, 13751
 ; OTHER INFORMATION: n = c or a
 ; NAME/KEY: misc_feature
 ; LOCATION: 716, 1293, 2401, 2429, 2618, 3083, 3125, 3635, 4256, 4898,
 ; LOCATION: 5062, 5167, 11069, 1398, 14479, 14730, 14796, 15344, 15450,
 ; LOCATION: 15503, 15590, 15840, 16149
 ; OTHER INFORMATION: n = a or g
 ; NAME/KEY: misc_feature
 ; LOCATION: 732, 1379, 1590, 2488, 3212, 5006, 11238, 11422, 11686,
 ; LOCATION: 12598, 13171, 13645, 13782, 13806, 13813, 14586, 14788,
 ; LOCATION: 15042, 15546, 15770
 ; OTHER INFORMATION: n = c or t

Qy 1021 TACTGGGAGACATCCACATCCTCCGCTGGCGGTGAGAGAGCAGCCGACGGGC 1080
Db 1097 TACTGGGAGACATCCACATCCTCCGCTGGCGGTGAGAGAGCAGCCGACGGGC 1156
Qy 1081 ACCGGCACCTCTCGGTATTCGGCATGACGGCGCGTACAGCCTGAGAGGTGACGGCGG 1140
Db 1157 ACCGGCACCTCTCGGTATTCGGCATGACGGCGCGTACAGCCTGAGAGGTGACGGCGG 1216
Qy 1141 GGCCCTCGGGAGCGGGTGGCGGAGGAGGAGGAGGCGCGCTGGGGG 1187
Db 1217 GGCCCTCGGGAGCGGGTGGCGGAGGAGGAGGAGGCGCGCTGGGGG 1263

RESULT 14

US-09-963-333-8

Sequence 8, Application US/09963333

GENERAL INFORMATION:

APPLICANT: Stanton, Jr., Vincent P.

TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES

TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT

FILE REFERENCE: 11926-015002

CURRENT APPLICATION NUMBER: US/09/963,333

CURRENT FILING DATE: 2001-09-24

PRIOR APPLICATION NUMBER: 09/658,659

PRIOR FILING DATE: 2000-09-08

PRIOR APPLICATION NUMBER: 09/596,033

PRIOR FILING DATE: 2000-06-15

PRIOR APPLICATION NUMBER: 09/357,743

PRIOR FILING DATE: 1999-07-20

PRIOR APPLICATION NUMBER: 09/357,024

PRIOR FILING DATE: 1999-07-19

PRIOR APPLICATION NUMBER: 60/093,484

PRIOR FILING DATE: 1998-07-20

NUMBER OF SEQ ID NOS: 16

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 8

LENGTH: 18597

TYPE: DNA

ORGANISM: Homo sapiens

NAME/KEY: misc_feature

LOCATION: 701, 13751

OTHER INFORMATION: n = c or a

NAME/KEY: misc_feature

LOCATION: 716, 1293, 2401, 2429, 2618, 3083, 3125, 3635, 4256, 4898,

LOCATION: 5062, 5167, 11069, 13298, 14479, 14730, 14796, 15344, 15450,

LOCATION: 15503, 15590, 15840, 16149

OTHER INFORMATION: n = a or g

NAME/KEY: misc_feature

LOCATION: 732, 1379, 1590, 2488, 3212, 5006, 11238, 11422, 11686,

LOCATION: 12598, 13171, 13645, 13782, 13806, 13813, 14586, 14788,

LOCATION: 15042, 15546, 15770

OTHER INFORMATION: n = c or t

NAME/KEY: misc_feature

LOCATION: 1322, 1688

OTHER INFORMATION: n = c or g

NAME/KEY: misc_feature

LOCATION: 2594, 11293, 16199, 16203

OTHER INFORMATION: n = g or t

NAME/KEY: misc_feature

LOCATION: 3619

OTHER INFORMATION: n = a or t

NAME/KEY: misc_feature

LOCATION: 14547

OTHER INFORMATION: nucleotide in position 14547 is t, or absent

US-09-963-333-8

Query Match 98.3%; Score 1167; DB 36; Length 18597;

Best Local Similarity 99.58; Pred. No. 4.5e-116;

Matches 1181; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Db 1157 ACCGGCACCTTCGGTATTCGGCANGCAGGCGCGCTACAGCTGAGAGTGACGGCGG 1216
Qy 1141 GGCCCTCGGGAACGGGTGGGGAAGGAGGAGGCGGCTGGGGA 1187
Db 1217 GGCCCTCGGGAACGGGTGGGGAAGGAGGAGGCGGCTGGGGA 1263

RESULT 15

US-09-652-109-9676
; Sequence 9676, Application US/09652109
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1180-001
; CURRENT APPLICATION NUMBER: US/09/652,109
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,128
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 10105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9676
; LENGTH: 1829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-109-9676

Query Match 28.2%; Score 335; DB 25; Length 1829;
Best Local Similarity 98.5%; Pred. No. 3.7e-27;
Matches 338; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 791 CGGCCGCGGAAAGCGCGGAGGGGTCCTGCCACCGCGCCACTTGGCTGCTCCG 850
Db 6 CGTCCGCGGAAAGCGCGGAGGGGTCCTGCCACCGCCACTTGGCTGCTCCG 65
Qy 851 TCCGCGCGCGCCACTTGGCTGCTCCGTCGCCCGCGCCACTTGGCTGCTCCG 910
Db 66 TCCGCGCGCGCCACTTGGCTGCTCCGTCGCCCGCGCCACTTGGCTGCTCCG 125
Qy 911 CCGCCCGCGCGCCACTTGGCTGCTCCGTCGCCCGCGCCACTTGGCTGCTCCG 970
Db 126 CCGCCCGCGCGCCACTTGGCTGCTCCGTCGCCCGCGCCACTTGGCTGCTCCG 185
Qy 971 CCGCACAGGAGCGGAGCGCGAGCGCGTCGCCCGCGCCACTTGGCTGCTCCG 1030
Db 186 CCGCACAGGAGCGGAGCGCGAGCGCGTCGCCCGCGCCACTTGGCTGCTCCG 245
Qy 1031 AGATCCACACATCTCCGTCGCCCGGCTCAGGAAGGAGACCGCGCCACCGGCA 1090
Db 246 AGATCCACACATCTCCGTCGCCCGGCTCAGGAAGGAGACCGCGCCACCGGCA 305
Qy 1091 TGTCGGTATTCGGCATGCGGCGCGCTACAGCTGAGAGTGA 1133
Db 306 TGTCGGTATTCGGCATGCGGCGCGCTACAGCTGAGAGTGA 348

Search completed: November 29, 2002, 00:16:39
Job time : 2476.31 secs

ib	677	TTCTCCAAAGGGCGCAGTCTTCTCCAGCCACCCGCACTGTGATCCAGGTTCCTCGGGTTCTCT	736
iy	661	AAGACTCTCAGGTGTGGCCCTCGGGCTCCGTCTCTGTGACCAACCCGTGGCTCTCGTGGTTTC	720
ib	737	AAGACTCTCAGGTGTGGCCCTCGGGCTCTGTGACCAACCCGTGGCTCTCGTGGTTTC	796
iy	721	CCCTCTGGCGCACGTCTCTAGAGCGGGGGCCGCGCGACCCCGCCGACGAGGAAGAAGCGG	780
ib	797	CCCTCTGGCGCACGTCTCTAGAGCGGGGGCCGCGCGACCCCGCCGACGAGGAAGAAGCGG	856
iy	781	GAGCGGGGACGGCCCGGGGAAAGGCGCGCGGAAGGGGTCTTGCACACCGCGCACTTTGG	840
ib	857	GAGCGCGGACGGCCCGGGGAAAGGCGCGCGGAAGGGGTCTTGCACACCGCGCACTTTGG	916
iy	841	CTTGCTCTCGTCCCGCCGCGCGCACCTTGGCGTCTCGTCCCGCCGCGCGCACCTTGGCGTCTG	900
ib	917	CTTGCTCTCGTCCCGCCGCGCGCACCTTGGCGTCTCGTCCCGCCGCGCGCACCTTGGCGTCTG	976
iy	901	CTTCTCGTCCCGCCGCGCGCGCACCTTGGCGTCTCGTCCCGCCGCGCGCACCTTGGCGTCTG	960
ib	977	CTTCTCGTCCCGCCGCGCGCGCACCTTGGCGTCTCGTCCCGCCGCGCGCACCTTGGCGTCTG	1036
iy	961	TTGCCCCCGCGCACAGGAGCGGAGCGCCGACGCGCTCCGCGCGCACGCGGGAGCTTCGAG	1020
ib	1037	TTGCCCCCGCGCACAGGAGCGGAGCGCCGACGCGCTCCGCGCGCACGCGGGAGCTTCGAG	1096
iy	1021	TACTGGGGCAGATCCAAACATCTCCGCTGCGGGCTCAGGAAGGACGACCGCACCGGCG	1080
ib	1097	TACTGGGGCAGATCCAAACATCTCCGCTGCGGGCTCAGGAAGGAGCGACCGCACCGGCG	1156
iy	1081	ACCGGCAACCTGTCTGGTATTTCGGCATGCAAGCGCGCTTACAGGCTTGAGAGTGAAGCGCG	1140
ib	1157	ACCGGCAACCTGTCTGGTATTTCGGCATGCAAGCGCGCTTACAGGCTTGAGAGTGAAGCGCG	1216
iy	1141	GGCCCCCTCGGGGAAGGGTGTGGGGGAAGGAGGGAGCGCGGTGGGGGA	1187
ib	1217	GGCCCCCTCGGGGAAGGGTGTGGGGGAAGGAGGGAGCGCGGTGGGGGA	1263

```

RESULT 3
JS-09-880-107-1590
; Sequence 1590, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherif, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1590
; LENGTH: 18596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D00596
US-09-880-107-1590

Query Match      98.3%; Score 1167; DB 10; Length 18596;
Best Local Similarity 99.2%; Pred. No. 1.2e-176;
Matches 1179; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY      1  GATCGCCCACTGCATCTCAGCTTGGGTGAGAGCGCAGCTTGTCTTCAAAAAAAAAAAA 60

```

Db	78	GATGCGCCACTGCGACTCTCCAGCCTGGGTGAGAGGCGAGACTCTGTCTCAAAAATAAAAAA	137
Qy	61	AAAAAGACCGCCAGGGCTCAAAACAAAAACCTCGAAAGACCTGGCGGTCTTTTTTTTT	120
Db	138	AAAAAGACCGCCAGGGCTCAAAACAAAAACCTCGAAAGACCTGGCGGTCTTTTTTTTT	197
Qy	121	TTTTTTTTTTTTTTTTTTTGGGACAGTCTTGCTCTGTGCGCCAGGCTGAGTACAAATG	180
Db	198	TTTTTTTTTTTTTTTTTTTGGGACAGTCTTGCTCTGTGCGCCAGGCTGAGTACAAATG	257
Qy	181	CGGATCTTGGCTCACTGCAACCTCTGCTCCCAAGGTTCAAGCAATTTCTTGCTCTCAGCC	240
Db	258	CGGATCTTGGCTCACTGCAACCTCTGCTCCCAAGGTTCAAGCAATTTCTTGCTCTCAGCC	317
Qy	241	TCCCAAGTAGCCACCAAGCCCACTAAATTTTTGTANTTTTAGTAGAGCGGGGTTTCA	300
Db	318	TCCCAAGTAGCCACCAAGCCCACTAAATTTTTGTANTTTTAGTAGAGCGGGGTTTCA	377
Qy	301	CATGTTCTCAGGCTGGTCTTGAACTCTGACCTCAGGTTGATCCACCGCTCGGCCCCC	360
Db	378	CATGTTCTCAGGCTGGTCTTGAACTCTGACCTCAGGTTGATCCACCGCTCGGCCCCC	437
Qy	361	CAAAAGTACTAGGANTTACAGGCTGAGCCACCGCGTCCAGCGCCTTGCGGTCTTTTAATCA	420
Db	438	CAAAAGTACTAGGANTTACAGGCTGAGCCACCGCGTCCAGCGCCTTGCGGTCTTTTAATCA	497
Qy	421	AGTAGAAAAGCTGCATATATACATCTGTCTTCTGTTGCTGNTTCAAGTAGAGAAATGG	480
Db	498	AGTAGAAAAGCTGCATATATACATCTGTCTTCTGTTGCTGNTTCAAGTAGAGAAATGG	556
Qy	481	AAATGCAAAATCCTTATTTAGTTCTTAGGAAAACAGATCTCAACACGACGTTTTGTGACAG	540
Db	557	AAATGCAAAATCCTTATTTAGTTCTTAGGAAAACAGATCTCAACACGACGTTTTGTGACAG	616
Qy	541	ACCGCAGAAAACTGGGAACTGTGCTGTGGCTTAGAGAGGCGCGCTCGACAGACGG	600
Db	617	ACCGCAGAAAACTGGGAACTGTGCTGTGGCTTAGAGAGGCGCGCTCGACAGACGG	676
Qy	601	TTTCCAAAGGGCGAGTCTTCTCCGACCAACGCACTGCTGNTTCAAGTTCCCGGGTTCCT	650
Db	677	TTTCCAAAGGGCGAGTCTTCTCCGACCAACGCACTGCTGNTTCAAGTTCCCGGGTTCCT	736
Qy	661	AAGACTCTCAGCTGTGGCCCTGGGCTCGGTCTGTGTGGCAACCGTGGCTCTGGTTC	720
Db	737	AAGACTCTCAGCTGTGGCCCTGGGCTCGGTCTGTGTGGCAACCGTGGCTCTGGTTC	796
Qy	721	CCCTTGGCGCAGCTCTCTAGAGCGGGGGCGCGCGACCCCGCGAGCAGGAGAGAGCG	780
Db	797	CCCTTGGCGCAGCTCTCTAGAGCGGGGGCGCGCGACCCCGCGAGCAGGAGAGAGCG	856
Qy	781	GAGCGCGGACGCGCGGGAAAAGCGCGCGGAGGGGTCTCTGCCACCGCGCTCTGG	840
Db	857	GAGCGCGGACGCGCGGGAAAAGCGCGCGGAGGGGTCTCTGCCACCGCGCTCTGG	916
Qy	841	CCTGCTCTCGTCTCCGCGCGCACTTGGAGCTGCTCTGCTCCGCGCGCGCACTTGGCTG	900
Db	917	CCTGCTCTCGTCTCCGCGCGCACTTGGAGCTGCTCTGCTCCGCGCGCGCACTTGGCTG	976
Qy	901	CCTCGGTCCTCCGCGCGCGCACTGCTGTGCGGGCTGCGAGCTGCGCGCGCGGCC	960
Db	977	CCTCGGTCCTCCGCGCGCGCACTGCTGTGCGGGCTGCGAGCTGCGCGCGCGGCC	1036
Qy	961	TTTGCCCCCGCGCGCAGAGCGGGACCGCGAGCGCGTCTCCGCGCAGCGGGAGTGCAG	1020
Db	1037	TTTGCCCCCGCGCGCAGAGCGGGACCGCGAGCGCGTCTCCGCGCAGCGGGAGTGCAG	1096
Qy	1021	TACCTGGGCGAGTCCAAACATCTCTCGCTGCGGCTGAGGAAGACGACGCGCAGGGC	1080
Db	1097	TACCTGGGCGAGTCCAAACATCTCTCGCTGCGGCTGAGGAAGACGACGCGCAGGGC	1156
Qy	1081	ACGGCACCTCTGTGGTATTCGCAATGACGGCGCGCTAAGCTTGAAGGTATCGCCG	1140
Db	1157	ACGGCACCTCTGTGGTATTCGCAATGACGGCGCGCTAAGCTTGAAGGTATCGCCG	1216

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

DN nucleic - nucleic search, using sw model

run on: November 28, 2002, 19:32:59 ; Search time 48.8226 Seconds
(without alignments)

7456.079 Million cell updates/sec

Title: US-09-963-333-7

Perfect score: 1187

Sequence: 1 gatcgccactgcactcca.....gaggaggcggtgggga 1187

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/6C.COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1167	98.3	18596	4	US-09-318-448-11
2	299.4	25.2	1536	3	US-09-089-195-1
3	299.4	25.2	1536	4	US-09-367-007C-38
4	207.4	17.5	942	4	US-09-347-878-29
5	198.6	16.7	40000	4	US-09-780-049-18
6	197.4	16.6	45546	4	US-09-146-053-6
7	193.2	16.3	111282	4	US-09-754-250-3
8	192	16.2	29629	4	US-09-729-995-3
9	191.8	16.2	14747	4	US-09-608-285A-42
10	191.8	16.2	15977	4	US-09-608-285A-59
11	190.6	16.1	162450	4	US-09-345-882-1
12	189.4	16.0	3804	2	US-08-483-488-5
13	189	15.9	631	4	US-09-385-982-354
14	189	15.9	3867	4	US-09-347-114A-81
15	189	15.9	35060	3	US-08-814-095-7
16	188.8	15.9	2713	2	US-08-916-901-6
17	188.8	15.9	2713	4	US-09-154-602-6
18	188.8	15.9	87350	3	US-08-781-891-79
19	188.8	15.9	87543	4	US-09-791-211-3
20	188.2	15.9	62804	4	US-09-800-960-3
21	187.6	15.8	17327	1	US-07-906-871-15
22	187.2	15.8	72604	4	US-09-268-992-7
23	187.2	15.8	72604	4	US-09-657-474-7
24	186.6	15.7	1019	4	US-09-177-650-128
25	186.2	15.7	32042	4	US-09-245-281-43
26	186	15.7	8453	4	US-09-167-681-45
27	185.8	15.7	489	4	US-09-370-838-109

US-09-318-448-11
; Sequence 11, Application US/09318448
; Patent No. 6210350
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Steiros, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; TITLE OF INVENTION: DEVELOPMENTAL DISORDERS
; FILE REFERENCE: 601-1-057
; CURRENT APPLICATION NUMBER: US/09/318,448
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 11
; LENGTH: 18596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-318-448-11

Query Match 98.3%; Score 1167; DB 4; Length 18596;
Best Local Similarity 99.2%; Pred. No. 1.5e-207;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Qy	1	GATCGGCGCAGCTGCACTCCAGCTGGGTGAGAGAGGAGAGACTCTGTCTCAAAAAA	60
Db	78	GATCGGCGCAGCTGCACTCCAGCTGGGTGAGAGAGGAGAGACTCTGTCTCAAAAAA	137
Qy	61	AAAAAGACCGCAGGCTCAAAACAAAAACCTCGAAAAAGCCCTGGCGGTCTTTT	120
Db	138	AAAAAGACCGCAGGCTCAAAACAAAAACCTCGAAAAAGCCCTGGCGGTCTTTT	197
Qy	121	TTTTTTTTTTTTTTTTTTCGGACAGTCTGTCTGTGCGCCAGGTGAGTACATG	180
Db	198	TTTTTTTTTTTTTTTTTTCGGACAGTCTGTCTGTGCGCCAGGTGAGTACATG	257
Qy	181	CGGATCTTGGTCTACTGCAACCTTGCCTCCAGGTTCAGCAATTCCTTCGCTCAGCC	240
Db	258	CGGATCTTGGTCTACTGCAACCTTGCCTCCAGGTTCAGCAATTCCTTCGCTCAGCC	317
Qy	241	TCCCAAGTAGCAGCAGCCAGCTAAATTTTGTATTTTAGTAGAGCGGGGTTTCAC	300
Db	318	TCCCAAGTAGCAGCAGCCAGCTAAATTTTGTATTTTAGTAGAGCGGGGTTTCAC	377
Qy	301	CATGTTTCCAGGTGTCTGTAACCTCTGACCTCAGTGCATCCACCGCTCGGCCCCC	360
Db	378	CATGTTTCCAGGTGTCTGTAACCTCTGACCTCAGTGCATCCACCGCTCGGCCCCC	437
Qy	361	CAAGTACTAGGANTACAGCGGTGAGCCACCGCGCTCCAGCGCCCTGGCGGTTTTTA	420

ALIGNMENTS

28	185.8	15.7	6769	1	US-08-480-784-20	Sequence 20, Appl	
29	185.8	15.7	6769	1	US-08-483-553-20	Sequence 20, Appl	
30	185.8	15.7	6769	1	US-08-487-002-20	Sequence 20, Appl	
31	185.8	15.7	6769	1	US-08-483-554B-20	Sequence 20, Appl	
32	185.8	15.7	6769	1	US-08-488-011B-20	Sequence 20, Appl	
33	185.8	15.7	6769	1	US-08-850-727-20	Sequence 20, Appl	
34	185.8	15.7	6769	5	PCT-US95-10203-20	Sequence 20, Appl	
35	185.8	15.7	6769	5	PCT-US95-10203-20	Sequence 20, Appl	
36	185.8	15.7	6769	5	PCT-US95-10203-20	Sequence 20, Appl	
37	185.2	15.6	70000	4	US-09-851-896-3	Sequence 3, Appl	
38	184	15.5	36551	4	US-09-738-894A-3	Sequence 3, Appl	
c	39	183.8	15.5	1043	4	US-09-165-868-4	Sequence 358, App
40	183.6	15.5	619	4	US-09-385-982-358	Sequence 3, Appl	
41	183.4	15.5	84495	4	US-09-797-906-3	Sequence 1, Appl	
42	183	15.4	168575	4	US-09-426-290-1	Sequence 1, Appl	
c	43	182.6	15.4	11288	4	US-08-646-301A-1	Sequence 1, Appl
c	44	182.6	15.4	11288	4	US-08-481-368A-4	Sequence 1, Appl
c	45	182.6	15.4	11288	4	US-08-154-712B-4	Sequence 4, Appl

OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
NAME/KEY: allele
LOCATION: 90813..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59

```

FEATURE:
; NAME/KEY: allele
; LOCATION: 108127..108177
; OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
; NAME/KEY: allele
; LOCATION: 108127..108177
; OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:
Query Match      16.1%; Score 190.6; DB 4; Length 162450;
Best Local Similarity 73.0%; Pred. No. 6.4e-27;
Matches 278; Conservative 0; Mismatches 83; Indels 20; Gaps
QY 116 TTTTITTTTTTTTTTTTTTTTTTTTTTGGAACAGCTTGCTCTGTGCCCAGGCTGGAGTACA 175
DB 72154 TTTTITTTTTTTTTTTTTTTTTTTTGTAGGCACAGACTTGCTGTCAACCAGGCTGGAGTACA 72213
QY 176 ATGGTCGGATCTGTGCTCACGTGAACCTCTGCCCTCCAGGTTCAAGCAATTCTCTGGCT 235
DB 72214 GTGGCGGATCTCGGCTCACTGCAACCTCCACTCCAGGCTCAAGCAATTCTCTGGCT 72273
QY 236 CAGCCTCCCAAGTA-----GCCACACGGCCAGCTAAATTTTGTAN 276
DB 72274 CAGCCTCCCGAGTAGCTGGGATTACAGGCGCTGCCACACGCTGCCCAATTTTGTAC 72333
QY 277 TTTTATGATGACGGGGGTTTCAACATGTGTCCAGGCTGCTCTNGAACCTCTCGACCTCA 336
DB 72334 TTTTGTGTAGAGAC-GGGGTTTTCACATGTTGGCAGGCTGCTCTCGAATCTCTGACCTCA 72392
QY 337 GGTGATCCACCGCGCTCGCGCCCCCAAAGTACTAGGATTA CAGCGTGTAGCCACCGGTC 396
DB 72393 AGTGATCCACTGCTCGACGCCCTCAAAGTGTGGATTGGAGTTG CAGGCGTGTAGCCACTGTGT 72452
QY 397 CAGCGCGCTGGCGGTTTTTAATCAAGTAGAAAAAGCTGCATATACCACTTGTCTTONGTTG 456
DB 72453 CAGCCCAACTGCTGGTACTTTTAAATGAAGTCAATAGTCAATTATTCTTCCACCATCAG 72512
QY 457 CNTTCACTGTAGAACGAGAAA 477
DB 72513 CACTCAGTTAAAGAAGAACAA 72533

RESULT 12
US-08-483-488-5
; Sequence 5, Application US/08483488
; Patent No. 5953985
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Johannes; Masters, Colin;
; APPLICANT: Beyreuther, Konrad
; TITLE OF INVENTION: Promoter of the Gene for the
; TITLE OF INVENTION: Human Precursor of the Alzheimer's
; TITLE OF INVENTION: Disease and its Use
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: SPRUNG HORN KRAMER & WOODS
; STREET: 660 White Plains Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10591-5144
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
; MEDIUM TYPE: storage
; COMPUTER: NEC Powermate SX/20
; OPERATING SYSTEM: DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,488
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/325,745

```

[illegible]


```
LOCATION: complement (33493..33591)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 4
FEATURE:
NAME/KEY: exon
LOCATION: complement (33297..33408)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 5
FEATURE:
NAME/KEY: exon
LOCATION: complement (32959..33094)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: exon
LOCATION: complement (32569..32628)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 7
FEATURE:
NAME/KEY: exon
LOCATION: complement (32386..32468)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 8
FEATURE:
NAME/KEY: exon
LOCATION: complement (31894..32080)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 9
FEATURE:
NAME/KEY: exon
LOCATION: complement (31363..31534)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 10
FEATURE:
NAME/KEY: exon
LOCATION: complement (31131..31284)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 11
FEATURE:
NAME/KEY: exon
LOCATION: complement (30816..31011)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 12
FEATURE:
NAME/KEY: exon
LOCATION: complement (30470..30626)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 13
FEATURE:
NAME/KEY: exon
LOCATION: complement (30187..30274)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 14
FEATURE:
NAME/KEY: exon
LOCATION: complement (29945..30073)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 15
FEATURE:
NAME/KEY: exon
LOCATION: complement (29664..29856)
OTHER INFORMATION: /gene= "ARS"
OTHER INFORMATION: /number= 16
```

US-08-814-095-7

Query Match 15.9%; Score 189; DB 3; Length 35060;

Best Local Similarity 72.5%; Pred. No. 1e-26;
Matches 308; Conservative 0; Mismatches 92; Indels 25; Gaps 4;

```
QY 1 GATCGCCGCTCCTCAGCTGGTGGAGAGCGAGACTCTGTCTCAAAAAAAAAA 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8322 GATTATGCCCTACTACAGCTGGGTGAGAGAGAGACCTGTCCCTTAAAGAAA 8381
```

```
QY 61 AAAA---AGACCGCCAGGGCTCAAAACAAAAAAGCCTCGGAAAAAGCCCTGGCGGTCTTTTT 117
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8382 AAAATAGATTGGGAAAAGTAGATCTTTATAGAGCCGAACACCTTTTATTATTATT 8441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 118 TTTTTTTTTTTTTTTTTTTTGGGAC--AGTCTTCTCTGTGCCCCAGGCTGGAGTACA 175
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8442 TATTTATTATTATTATTATTATTTTGGAGACAGAGTCTGCTCTGTGCCCCAGGCTGGAGGCA 8501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 176 ATGTCGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTTCGCT 235
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8502 GTGGTCGATCTCGGCTCACTGCAACCTCTACCTCTAGATTCAAGCAATTTCTTCGCT 8561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 236 CAGCTCCCAAGTA-----GCCACACGCCCCAGCTAA-TTTTTGTA 275
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8562 CAGTCTCCCAATAGCTGGGATTACAGGCGCCGCCACACAGGCCAGATAATTTTGTGA 8621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 276 NTTTTAGTAGACGGGGTTTCAACCATGTTGTCCAGGCTGGTCTNGAACTCTGACCTC 335
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8622 TTTTATTAGTAGACAGGGTTTTCCTCATGTTGGCCAGGCTGGTCTTGAACCTCCTAACCTC 8681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 336 AGGTGATCCACCGCTCGGCCCCCAAAAGTACTAGGATTACAGCGCTGAGCCACCGCT 395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8682 AGGTGATCCACCGCTCGGCCCTCTCAAGTACTGGGATTACAGGATGAGCCACCAAC 8741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 396 CCAGC 400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8742 CCAGC 8746
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: November 28, 2002, 22:48:48

Job time : 311.823 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

M nucleic - nucleic search, using sw model

run on: November 28, 2002, 17:24:19 ; Search time 234.959 Seconds
(without alignments)
11376.979 Million cell updates/sec

Title: US-09-963-333-7
Perfect score: 1187
Sequence: 1 gatcgccactgcactcca.....ggggggcgcggtgggga 1187

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 101002:*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1167	98.3	18596	22 AAF31109	Thymidylate synthase
2	1167	98.3	18596	22 AAC91215	Human thymidylate
3	1167	98.3	18596	24 AEN95092	Gene #1590 used to
4	1167	98.3	18596	24 ABK43334	Human thymidylate
5	1167	98.3	18596	24 ABL62854	Breast cancer rela
6	1167	98.3	18596	24 ABL63078	Breast cancer rela
7	1167	98.3	18596	24 ABL67927	Ovary cancer relat
8	770.6	64.9	1186	24 ABK40090	Human chemically p
9	710	59.8	1186	24 ABK40089	Human chemically p

10	299.4	25.2	1536	24	ABK43335	Human Thymidylate
11	288.8	24.3	3298	24	AAS94945	Human DNA sequence
12	226.6	19.1	764	24	ABQ39526	Oligonucleotide fo
13	226.6	19.1	764	24	ABQ39527	Oligonucleotide fo
14	223.2	18.8	764	24	ABQ39528	Oligonucleotide fo
15	223.2	18.8	764	24	ABQ39529	Oligonucleotide fo
16	216.2	18.2	136284	24	ABK83575	Human cDNA differe
17	207.4	17.5	1539	23	AAS94960	DNA encoding novel
18	206.8	17.4	9519	22	AAI37169	Human musculoskele
19	206.8	17.4	9519	23	ABK42169	Genomic sequence #
20	206.8	17.4	56632	22	AAK65581	Human immune/haema
21	204.8	17.3	9731	22	ABL04943	Human reproductive
22	204.8	17.3	9731	23	ABL97837	Human testicular a
23	204.6	17.2	11453	22	AAK73272	Human immune/haema
24	204.6	17.2	11553	22	AAK73271	Human immune/haema
25	204.2	17.2	37449	22	AAK66874	Human immune/haema
26	204.2	17.2	56743	22	AAK68202	Human immune/haema
27	204.2	17.2	56743	22	AAK81760	Human immune/haema
28	203.6	17.2	20601	22	AAK79760	Human immune/haema
29	202.8	17.1	5591	22	ABAI5206	Human nervous syst
30	202.8	17.1	5591	22	AAS29200	Genomic sequence #
31	202.6	17.1	22916	22	AAK65305	Human immune/haema
32	201.8	17.0	10820	22	ABAI6598	Human nervous syst
33	201.8	17.0	122888	24	ABK83569	Human cDNA differe
34	200	16.8	9060	24	AAS19272	Human genomic DNA
35	200	16.8	9060	24	AAS19294	Human generic DNA
36	199.4	16.8	32169	22	ABAI4358	Human nervous syst
37	198.6	16.7	32173	22	ABA07767	Human ovarian and
38	198.6	16.7	32173	22	AAI03558	Human reproductive
39	198.6	16.7	38928	22	AAK68452	Human immune/haema
40	198.6	16.7	109906	24	ABK94411	DNA encoding endot
41	197.8	16.7	846	22	AAS44955	CDNA encoding nove
42	197.8	16.7	20050	24	ABL53875	Human immunoglobul
43	197.6	16.6	7017	22	AAI37408	Human musculoskele
44	197.6	16.6	50000	21	AAA96365	Polymorphic repeat
45	197.4	16.6	45546	20	AAK23520	Human kidney amino

ALIGNMENTS

RESULT 1
AAF31109
ID AAF31109 standard; cDNA; 18596 BP.
XX
AC AAF31109;
XX
DT 27-APR-2001 (first entry)
XX
DE Thymidylate synthase coding sequence.
XX
KW Analyte-binding enzyme; analyte analysis; ss.
XX
OS Homo sapiens.
XX
PN WO200102600-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18057.
XX
PR 06-JUL-1999; 99US-0347878.
PR 06-DEC-1999; 99US-0457205.
XX
(GEAT) GEN ATOMICS.
XX
PI Yuan C;
XX
DR WPI; 2001-071583/08.
XX
PT Assaying method, useful for prognosis and diagnosis of disease,
PT comprises contacting sample with a mutant analyte-binding enzyme and
PT detecting binding -

X	Sequence	18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;
Q	Query Match	98.3%; Score 1167; DB 22; Length 18596;
	Best Local Similarity	99.2%; Pred. No. 1.1e-193;
	Matches 1170; Conservative	0; Mismatches 8; Indels 1; Gaps
y	1	GATCGCGCACCTGCCTCCAGCCTGGGTGAGAGACGAGACTCTGTCTCAAAAAAAA 60
b	78	GATCGCGCACCTGCCTCCAGCCTGGGTGAGAGACGAGACTCTGTCTCAAAAAAAA 137
y	61	AAAAAGACCGCCAGGGCTCAAAACAAAAACCTCGAAAGCCCTCGGCGGTCTTTTTTTTT 120
b	138	AAAAAGACCGCCAGGGCTCAAAACAAAAACCTCGGAAAGCCCTGGCGGTCTTTTTTTTT 197
y	121	TTTTTTTTTTTTTTTTTTTTTTGGGACAGTCTTGTCTGTCCGCGAGGTGAGTACAAATGGT 180
b	198	TTTTTTTTTTTTTTTTTTTTTTGGGACAGTCTTGTCTGTCCGCGAGGTGAGTACAAATGGT 257
y	181	CGGATCTTGGCTCACTGCACAACTCTGCTCTCCAGGTTCAAGCAATTTCTTCTGCTCAGCC 240
b	258	CGGATCTTGGCTCACTGCACAACTCTGCTCTCCAGGTTCAAGCAATTTCTTCTGCTCAGCC 317
y	241	TCCCAAGTAGCCACCAACGCCGACGTAATTTTGTGTANTTTTGTAGAGACGGGGTTTCAC 300
b	318	TCCCAAGTAGCCACCAACGCCGACGTAATTTTGTGTANTTTTGTAGAGACGGGGTTTCAC 377
y	301	CAGTGTGTCGAGCGTGTCTGAACTCTTGACCTCAGGTGATCCACCCGCTTCGGCCCCC 360
b	378	CATGTTGTCCAGCGTGGTCTTGAACCTCTGACCTCAGGTGATCCACCCGCTTCGGCCCCC 437
y	361	CAAAGTACTAGGATTTACAGGCGGTGAGCCACCGCGTCCAGCGCTCCGCGGTTTTTAATCA 420
b	438	CAAAGTACTAGGATTTACAGGCGGTGAGCCACCGCGTCCAGCGCTCCGCGGTTTTTAATCA 497
y	421	AGTGAAAGAGCTCAATTATACCACTTGCTTCNGTTGNTTTCAGTGAGACGAGAAATGG 480
b	498	AGTGAAAGAGCTCAATTATACCACTTGCTTCGTTGCTTGC--TTCAAGTGAGACGAGAAATGG 556
y	481	AAATGCAAATCNCNTTATTAGTTGTAGGAAACAGATCTCAAAACAGCAGTTTTTGTGACAAG 540
b	557	AAATGCANAATCCCTTATTAGTTGTAGGAAACAGATCTCAAAACAGCAGTTTTTGTGACAAG 616
y	541	ACCGCAGGAAAAAGTGGGAACTGTGTCTGTGGCTTAGAGAGGCGCGGTGACACGACGG 600
b	617	ACCGCAGGAAAAAGTGGGAACTGTGTCTGTGGCTTAGAGAGGCGCGGTGACACGACGG 676
y	601	TTCCCAAGGGCGAGTCTTCCNGCCACCGCACCTGCTGNTCCAGGTTCCCGGNTTCCCT 660
b	677	TTCCCAAGGGCGAGTCTTCCAGCCACCGCACCTGCTGNTCCAGGTTCCCGGNTTCCCT 736
y	661	AGAGCTCTCAGCTGTGCGCTGGCTCCGTTCTGTGTCCACACCCGTGGCTCTCGGTTTC 720
b	737	AGAGCTCTCAGCTGTGCGCTGGCTCCGTTCTGTGTCCACACCCGTGGCTCTCGGTTTC 796
y	721	CCCTGGCGCAGCTCTCTTAGAGCGGGGCGCGCGAACCCCGCCGACGAGGAGAGCGG 780
b	797	CCCTGGCGCAGCTCTCTTAGAGCGGGGCGCGCGAACCCCGCCGACGAGGAGAGCGG 856
y	781	GAGCGCGGAGCGCGCGGGAAGGCGCGGAAAGGGGTCTGTGCACCGCGCCTCTTGG 840
b	857	GAGCGCGGAGCGCGCGGGAAGGCGCGGGAAGGGGTCTGTGCACCGCGCCTCTTGG 916
y	841	CTTGCCTTCGGTCCCGCGCGCACTTTGAGCTGTGCTTCGTTCCCGCGCGCACTTGGCTG 900
b	917	CTTGCCTTCGGTCCCGCGCGCACTTTGAGCTGTGCTTCGTTCCCGCGCGCACTTGGCTG 976
y	901	CTTTCGTTCCCGCGCGCGCACTTGTGCTGTGCGGTCTGGAGTCTCCGCGCGCGCC 960
b	977	CTTTCGTTCCCGCGCGCGCACTTGTGCTGTGCGGTCTGGAGTCTCCGCGCGCGCC 1036
y	961	TTGCGCCCCCGCGCACAAGAGAGCGGACGCGAGCCGGGTCTCCGCGCAGCGGGAGCTGCAG 1020

Db	1037	TTGGCCCCCGCCGACACAGAGAGGGGACGCCGAGCCGCGTCCGCCGCCACAGGGGAGCTGCGAG	1096
Qy	1021	TACCTGGGGCGAGATCAGAAACATCTCCGCTGCGCGCTCAGGAAGACACACGACCGACACGGCC	1080
Db	1097	TACCTGGGGCGAGATCCACACATCTCCGCTGCGCGCTCAGGAAGACACGACCGACACGGCC	1156
Qy	1081	ACGGCAGCCCTGTCCGCTATTCCGCATCTCAGGCGCGCTACAGCTGAGAGGTTGACGCCGCC	1140
Db	1157	ACGGCAGCCCTGTCCGCTATTCCGCATCTCAGGCGCGCTACAGCTGAGAGGTTGACGCCGCC	1216
Qy	1141	GGCCCTTCGGGACCGGTGGCGGGAGAGAGGAGCGCGCTGGGGA	1187
Db	1217	GGCTTCGCGGACCGGTGGCGGGAGAGAGGAGCGCGCTGGGGA	1263

RESULT 3

ABN95092
ID ABN95092 standard: DNA: 18596 BP.

AC ABN95092:

13-AUG-2002 (first entry)

Conc #1590 used to diagnose liver cancer.

XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
metastatic liver tumour; cytostatic; expression profile; disease state;
KW drug toxicity; drug efficacy; drug metabolism.

AA Homo sapiens.

XX PN WO200229103-A2

XX
11-APR-2002

XX 02-OCT-2001: 2001WO-US30589.
 PF

XX
03-OCT-2000. 2000US-237054P.

XX
BA (GENE-) GENE LOGIC INC.

XX
Dr
XX

Alvares C.
Peres-Da-Silva S,
Vockley JG;

XX WPT: 2002-426119/45

xx
pt Diagnosing and detecting the progression of liver cancer,
pt hepatocellular carcinoma or metastatic liver tumor in a patient,
pt involves detecting the level of expression of two or more genes in a
pt liver tissue sample -
pt

XX
ps
Claim 1: SEO ID NO 1590; 298pp; English.

XX The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in A893503-A897455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.

CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at <http://www.int/pub/mib/published/pct/sequences>.

XX 19506 PD, 4521 A, 3991 C, 4479 G, 5605 T; 0 other;

Query Match 98.3%; Score 1167; DB 24; Length 18596;
Best Local Similarity 99.2%; Pred. No. 1.1e-193;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1081 ACCGCGACCCCTGTCGGTATTTCGGCATGCGAGCGCGCTACAGCTGAGAGGTGACCGCGG 1140
Db 1157 ACCGCGACCCCTGTCGGTATTTCGGCATGCGAGCGCGCTACAGCTGAGAGGTGACCGCGG 1216
QY 1141 GCGCCCTGCGGAGCGGTGCGGGAAGAGGAGCGCGCGCTGGGA 1187
Db 1217 GCGCCCTGCGGAGCGGTGCGGGAAGAGGAGCGCGCGCTGGGA 1263

RESULT 4
ABK43334
ID ABK43334 standard; DNA; 18596 BP.
AC ABK43334;
XX
XX
DT 05-JUN-2002 (first entry)
XX
DE Human Thymidylate synthase gene sequence.
XX
KW HKNG1; ds; gene; chromosome 18p; bipolar affective disorder; BAD;
KW severe bipolar affective (mood) disorder; BP-I; schizophrenia; TS;
KW Hong Kong new gene 1; antimanic; antidepressant; neuroleptic;
KW Thymidylate synthase.
XX
OS Homo sapiens.
XX
PN W0200210366-A2.
XX
PD 07-FEB-2002.
XX
PF 02-AUG-2001; 2001WO-US24417.
XX
PR 02-AUG-2000; 2000US-0631275.
PR 28-NOV-2000; 2000US-0722544.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX (REGC) UNIV CALIFORNIA.
XX
XX Chen H, Freimer NB, Novak T;
XX WPI; 2002-195962/25.
XX
XX New nucleic acid molecule Hong Kong New Gene 1 (HKNG1), useful for
FT screening for molecules which modulate HKNG1 expression for the
PT treatment of bipolar disorder and schizophrenia -
XX Disclosure; Fig 44; 367pp; English.

The invention relates to an isolated nucleic acid molecule comprising a
nucleotide sequence that encodes a Hong Kong New Gene (HKNG) 1 gene
product. The human gene for HKNG1 is located on chromosome 18p in
an area associated with bipolar affective disorder, BAD. Also
included are an expression vector comprising the nucleic acid, a
host cell expressing the nucleic acid, an anti-HKNG1 antibody, a method
of identifying modulators of HKNG1, and identifying an individual (at
risk of) having HKNG1-mediated disorder comprising detecting the presence
or absence of a polymorphism that correlates with an HKNG1 allele
associated with the disorder, where the presence of the polymorphism
indicates that the individual (is at risk of) having HKNG1-mediated
disorder. A (small molecule) compound which modulates (inhibits or
potentiates) expression of a HKNG1 gene or gene product in a human
individual is useful for the treatment of a HKNG1-mediated disorder
such as bipolar affective disorder (BAD), severe bipolar affective (mood)
disorder (BP-I) and schizophrenia. The present sequence is a
thymidylate synthase, TS, genomic DNA sequence. The gene for TS
overlaps that of HKNG1 and therefore TS may also be involved in the
diseases listed above.

Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;
SQ

QY 1 GATCGCCCACTGCACTCCAGCCTGCGGTGAGAGAGCGAGCTCTGTCTCAAAAAA 60
Db 78 GATCGCCCACTGCACTCCAGCCTGCGGTGAGAGAGCGAGCTCTGTCTCAAAAAA 137
QY 61 AAAAAGACCCGAGGCTCAAA CAAAAA CTTGGAAGACCTGCGGCTCTTTTTTTT 120
Db 138 AAAAAGACCCGAGGCTCAAA CAAAAA CTTGGAAGACCTGCGGCTCTTTTTTTT 197
QY 121 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 180
Db 198 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 257
QY 181 CGGATCTGGCTCACTGCAACTCTGCTCCAGGTTCAAGCAATTTCTTCTGCTCAGCC 240
Db 258 CGGATCTGGCTCACTGCAACTCTGCTCCAGGTTCAAGCAATTTCTTCTGCTCAGCC 317
QY 241 TCCCAAGTAGCACCACCGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGGTTTCA 300
Db 318 TCCCAAGTAGCACCACCGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGGTTTCA 377
QY 301 CATGTTGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 378 CATGTTGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 437
QY 361 CAAAGTACTAGGATTAAGCGGTGAGCAACCGGCTCAGCGCCCTGCGGGTTTAAATCA 420
Db 438 CAAAGTACTAGGATTAAGCGGTGAGCAACCGGCTCAGCGCCCTGCGGGTTTAAATCA 497
QY 421 AGTAGAAAGTGCATTAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 498 AGTAGAAAGTGCATTAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556
QY 481 AAATGCAATGCTTATTTAGTTGTAGAAACAGATCTCAACAGCAGTCTTCTGACAG 540
Db 557 AAATGCAATGCTTATTTAGTTGTAGAAACAGATCTCAACAGCAGTCTTCTGACAG 616
QY 541 ACCGAGAGAAACCTGGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 617 ACCGAGAGAAACCTGGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
QY 601 TTCCCAAGAGCGCAGTCTTCCCGCCACCGCCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 677 TTCCCAAGAGCGCAGTCTTCCCGCCACCGCCTGCTGCTGCTGCTGCTGCTGCTGCT 736
QY 661 AAGACTCTCAGTGTGGCCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 737 AAGACTCTCAGTGTGGCCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
QY 721 CCCCTGGCGCAGCTCTCTAGAGGGGGGCGCGGACCCCGCGAGCAGGAGAGGGCG 780
Db 797 CCCCTGGCGCAGCTCTCTAGAGGGGGGCGCGGACCCCGCGAGCAGGAGAGGGCG 856
QY 781 GAGCGGGAAGCGCGGGAAGAGCGCGGAGGGGTCTGCGACCGCGCCTGCTGCTGCTGCT 840
Db 857 GAGCGGGAAGCGCGGGAAGAGCGCGGAGGGGTCTGCGACCGCGCCTGCTGCTGCTGCT 916
QY 841 CTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 917 CTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
QY 901 CTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 977 CTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
QY 961 TTGCCCCCGCGCAGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGGAGCTGAG 1020
Db 1037 TTGCCCCCGCGCAGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGGAGCTGAG 1096
QY 1021 TACTTGGGCGAGATCAACACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1097 TACTTGGGCGAGATCAACACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156

Query Match 98.3%; Score 1167; DB 24; Length 18596;
Best Local Similarity 99.2%; Pred. No. 1.1e-193;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Y	1	GATCGCGGCACCTGCACTCCAGCCTGGGTGAGAGAGCAGAGACTCTGTCTCAAAAAAAAAA	60
b	78	GATCGCGGCACCTGCACTCCAGCCTGGGTGAGAGAGCAGAGACTCTGTCTCAAAAAAAAAA	137
Y	61	AAAAAGACGCGCAGGGCTCAAAACAAAAACCTCGAAAAAGCCTGCGCGTCTTTTTTTTT	120
b	138	AAAAAGACGCGCAGGGCTCAAAACAAAAACCTCGAAAAAGCCTGCGCGTCTTTTTTTTT	197
Y	121	TTTTTTTTTTTTTTTTTTTTTGGGACAGTCTTGCTCTGTCCGCCAGGCTGGAGTACAATG	180
b	198	TTTTTTTTTTTTTTTTTTTTTGGGACAGTCTTGCTCTGTCCGCCAGGCTGGAGTACAATG	257
Y	181	CGGATCTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATCTCTCTGCTCGGCC	240
b	258	CGGATCTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATCTCTCTGCTCGGCC	317
Y	241	TCCCAAGTAGCCACGCGCCAGCTAAATTTTTTGTANTTTTTAGTAGAGACGGGGTTTCAC	300
b	318	TCCCAAGTAGCCACGCGCCAGCTAAATTTTTTGTANTTTTTAGTAGAGACGGGGTTTCAC	377
Y	301	CATGTGTTCAGGCTGGTCTGAGAACTCTCTGACCTCAAGTATCCACCGGCTCGGCCCCC	360
b	378	CATGTGTTCAGGCTGGTCTGAGAACTCTCTGACCTCAAGTATCCACCGGCTCGGCCCCC	437
Y	361	CAAGTACTAGGATTCAGGCTGAGCCACGCGTCCAGCGCCTGGCGTTTAAATCA	420
b	438	CAAGTACTAGGATTCAGGCTGAGCCACGCGTCCAGCGCCTGGCGTTTAAATCA	497
Y	421	AGTACAAAAGCTGCAATATACCATCTGCTCNGTTCGNTTCAGTGAGAACGAGAAATGG	480
b	498	AGTACAAAAGCTGCAATATACCATCTGCTCNGTTCGNTTCAGTGAGAACGAGAAATGG	556
Y	481	AAATGCAAAATCCTTATAGTTGTAGGAAAACAGATCTCAAACAGCAGTTTGTNGACAAG	540
b	557	AAATGCAAAATCCTTATAGTTGTAGGAAAACAGATCTCAAACAGCAGTTTGTNGACAAG	616
Y	541	ACCGCAGGAAAAGTGGGAACTGCTGCTGCTTGTAGAGAGGCGCGGTGACACAGCG	600
b	617	ACCGCAGGAAAAGTGGGAACTGCTGCTGCTTGTAGAGAGGCGCGGTGACACAGCG	676
Y	601	TTCCCAAAGGCGCAGTCTCTCCGNGCCACGCACTGCTCCAGTTCCCGGTTTCCT	660
b	677	TTCCCAAAGGCGCAGTCTCTCCGNGCCACGCACTGCTCCAGTTCCCGGTTTCCT	736
Y	661	AAGACTCTAGCTGTGGGCTCGGCTCTGTGTGCAACACCGCTGCTCTCGTTTC	720
b	737	AAGACTCTAGCTGTGGGCTCGGCTCTGTGTGCAACACCGCTGCTCTCGTTTC	796
Y	721	CCCTGCGGCACAGCTCTTAGAGCGGGGGCGCGGACCCCGCGAGCAGGAGAGGG	780
b	797	CCCTGCGGCACAGCTCTTAGAGCGGGGGCGCGGACCCCGCGAGCAGGAGAGGG	856
Y	781	GAGCGCGGACGCGCGCGGAAAAGGCGCGCAAGGGTCTCTGCCACCGCGCACTTGG	840
b	857	GAGCGCGGACGCGCGCGGAAAAGGCGCGCAAGGGTCTCTGCCACCGCGCACTTGG	916
Y	841	CTGCTCTCGTCCGCGCGCACTTGCGCTGCTCTCCGTCGCGCGCCACTTCGCTG	900
b	917	CTGCTCTCGTCCGCGCGCACTTGCGCTGCTCTCCGTCGCGCGCCACTTCGCTG	976
Y	901	CCTCGCTCCGCGCGCGCGCCATGCTCTGCGCGGCTCGGAGCTGCGCGCGCGCC	960
b	977	CCTCGCTCCGCGCGCGCGCCATGCTCTGCGCGGCTCGGAGCTGCGCGCGCGCC	1036
Y	961	TTGCGCCCGCGCGCACAGAGCGGAGCGCGAGCCGCGTCCGCGCAGCGGAGCTGCAG	1020
b	1037	TTGCGCCCGCGCGCACAGAGCGGAGCGCGAGCCGCGTCCGCGCAGCGGAGCTGCAG	1096
Y	1021	TACCTGGGGCAGATCCAAACATCTCCGCTGCGGGTTCAGGAAGGACGACGACCGGCG	1080
b	1097	TACCTGGGGCAGATCCAAACATCTCCGCTGCGGGTTCAGGAAGGACGACGACCGGCG	1156

PR 01-NOV-2000; 2000US-245084P.
XX (AVAL-) AVALON PHARM.
XX PA
XX PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
XX Claim 1; SEQ ID 1191; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX
SQ Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;
Query Match 98.3%; Score 1167; DB 24; Length 18596;
Best Local Similarity 99.2%; Pred. No. 1.1e-193;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 1 GATCGCGCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 60
DB 78 GATCGCGCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 137
QY 61 AAAAGACCGCAGCGCTCAACAAAAACCTCGGAAAGCCCTGGGGTCTTTTTTT 120
DB 138 AAAAGACCGCAGCGCTCAACAAAAACCTCGGAAAGCCCTGGGGTCTTTTTTT 197
QY 121 TTTTCTTTTTTTTTTTTTTTGGGACAGTCTGTCTGTCTGCTCCAGGCTGGAGTACATGGT 180
DB 198 TTTTCTTTTTTTTTTTTTTTGGGACAGTCTGTCTGTCTGCTCCAGGCTGGAGTACATGGT 257
QY 181 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGCTTCAAGCAATCTCTGCTCAGCC 240
DB 258 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGCTTCAAGCAATCTCTGCTCAGCC 317
QY 241 TCCCAAGTAGCCACAGCCCGCAGTAAATTTTGTANTTTTGTAGAGACGGGGTTTTCAC 300
DB 318 TCCCAAGTAGCCACAGCCCGCAGTAAATTTTGTACTTTTGTAGAGACGGGGTTTTCAC 377
QY 301 CATGTGTCCAGGCTGGTCTGCTGACCTCAGTGTGATCCACCCCGCTCGGCGCCCC 360
DB 378 CATGTGTCCAGGCTGGTCTGGAACCTCTGACTCAGTGTGATCCACCCCGCTCGGCCCC 437
QY 361 CAAAGTACTAGGATTTACAGGCTGAGCCACCGCTCCAGCGCTCGGCGTTTAAATCA 420
DB 438 CAAAGTACTAGGATTTACAGGCTGAGCCACCGCTCCAGCGCTCGGCGTTTAAATCA 497
QY 421 AGTAGAAAAAGCTGCATTTATACCACTTGTCTTCTGTTTCTGNTTTCAGTGAGAACGAAATGG 480
DB 498 AGTAGAAAAAGCTGCATTTATACCACTTGTCTTCTGTTTCTGNTTTCAGTGAGAACGAAATGG 556
QY 481 AAATGCAATCNCTATTATTAGTTGAGGAAACAGATCTCAACACAGCTTTTGTGACAAAG 540

DB 557 AAATGCAATCCCTATTATTAGTTGAGGAAACAGATCTCAACACAGCACTTTTGTGACAAAG 616
QY 541 ACGCAGAGAAACGTTGGGAACGTGCTGCTGCTTGTAGAGAGCGCGGTGCGACAGAGCG 600
DB 617 ACCGAGAGAAACGTTGGGAACGTGCTGCTGCTTGTAGAGAGCGCGGTGCGACAGAGCG 676
QY 601 TTCCAAAAGGCGCAGTCCCTTCCNGCCACACGCACTGCTGCTTCCAGGTTCCCGGGTTTCT 660
DB 677 TTCCAAAAGGCGCAGTCCCTTCCAGCACCGCACTGCTGCTTCCAGGTTCCCGGGTTTCT 736
QY 661 AAGACTCTCAGCTGTGGCCCTTGGGCTTCCGTTCTGTGCTCACACCGCTGGCTTCTCGGTTTC 720
DB 737 AAGACTCTCAGCTGTGGCCCTTGGGCTTCCGTTCTGTGCTCACACCGCTGGCTTCTCGGTTTC 796
QY 721 CCCCTGGCGCACTCTTCTTAGAGCGGGGCGCCGCGACCCCGCGAGAGAGAGAGCG 780
DB 797 CCCCTGGCGCACTCTCTTAGAGCGGGGCGCGCGGAGACCCCGCGAGAGAGAGCG 856
QY 781 GAGCGGGGAGCGCGCGGGGAAAGCGCGGAAAGGGGCTCTTGGCACCACCGCGCATTGG 840
DB 857 GAGCGCGGAGCGCGCGGGGAAAGCGCGGAAAGGGGCTCTTGGCACCACCGCGCATTGG 916
QY 841 CTTGGCTTCCGTCGCGCGCGCACTTGGGCTTCCGTTCCCGCGCGCACTTGGCCTG 900
DB 917 CTTGGCTTCCGTCGCGCGCGCACTTGGGCTTCCGTTCCCGCGCGCACTTGGCCTG 976
QY 901 CTTGGCTTCCGTCGCGCGCGCACTTGGGCTTCCGTTCCCGCGCGCACTTGGCCTG 960
DB 977 CTTGGCTTCCGTCGCGCGCGCACTTGGGCTTCCGTTCCCGCGCGCACTTGGCCTG 1036
QY 961 TTGCCCCCGCGCGCACAGGAGCGGAGCGCGCGCTCCGCGCGCACCGGGAGCTGCAG 1020
DB 1037 TTGCCCCCGCGCGCACAGGAGCGGAGCGCGCGCTCCGCGCGCACCGGGAGCTGCAG 1096
QY 1021 TACCTTGGGCGAGATCCAAACACATCTTCCGCTGCGCGGTTCAGGAGAGAGACCGCACCGG 1080
DB 1097 TACCTTGGGCGAGATCCAAACACATCTTCCGCTGCGCGGTTCAGGAGAGAGACCGCACCGG 1156
QY 1081 ACCGCACTCTGTGCTGATTTTGGCATGAGGCGCGCTACAGCTGAGAGTGAACCGCG 1140
DB 1157 ACCGCACTCTGTGCTGATTTTGGCATGAGGCGCGCTACAGCTGAGAGTGAACCGCG 1216
QY 1141 GGGCCCTCGGAGAGCGGTTGGGAGAGGAGGAGGAGCGCGCTGGGGA 1187
DB 1217 GGGCCCTCGGAGAGCGGTTGGGAGAGGAGGAGGAGGAGCGCGCTGGGGA 1263

RESULT 6
ABL63078 standard; DNA; 18596 BP.
ID ABL63078
XX
AC ABL63078;
XX
XX 15-MAY-2002 (first entry)
XX
XX Breast cancer related gene sequence SEQ ID NO:1415.
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
XX cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
XX gene; da.
OS Homo sapiens.
XX
XX W0200194629-A2.
XX
XX 13-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US10838.
XX
XX 05-JUN-2000; 2000US-209473P.
XX
XX 05-JUN-2000; 2000US-209531P.
XX
XX 18-SEP-2000; 2000US-233133P.
XX

QY 422 GTAGAAAGCTGATTATACACCTTGCTTCNGTTGCTTTCAGTGAGAACGAAAGTAATGGA 481
 Db 765 ATAAAAAACTACATATACACCTTACTTTCGATTAC-TTCAATAAAAACGAAATAATAA 707
 QY 482 AATGCATATCCTTATTAGTTGTAGGAAAACAGATCTCAACAGCAGTTTGTGTGACAGA 541
 Db 706 AATACAAATCCCTTATTATATATAAAAAACAATCTCAACACCAATTTTATTATAA 647
 QY 542 CCGCAGGAAACGTTGGAACTGTGTGCTGTCTAGAGAGCGCGTTCAGACGAGGT 601
 Db 646 CCGCAAAAAACGTAATACTATACTAATAAAAAAAGCGATCGACCAACGAT 587
 QY 602 TCCCAAGGCGAGTCTTTCGCGCAACGCGCTGNTCCAGTTCCTCCGGTTCCTTA 661
 Db 586 TCCCAAAAAACGTAATCTTCCCAACACCGCACTTACATCCCAATTTCCGAA 527
 QY 662 AGACTCTACGCTGTGGCTTCCGTTCTGTGCGCACACCGCTGCTCTGCGTTTC 721
 Db 526 AAATCTCACTATAAACCCTAACTCCGTTCTATACACACCCGTAATCTCTAGTTTC 467
 QY 722 CCTTGGCGCAGCTCTCTAGAGCGGGCGCGCGCACCCCGCAGCAGGAGAGCGG 781
 Db 466 CCTTAACGCGCTCTCTTAATAACGAAACCGCGCGCACCCCGCGACAAATAACGA 407
 QY 782 AGCGCGGAGCGCGCGGAAAGCGCGGAGAGGGGTCCTGCCACCGCGCACTTGGC 841
 Db 406 AACCGGAAACGACCGGAAAAAAGCGCGGAAAAATCTTACCACCGCGCACTTAAC 347
 QY 842 CTGCTCTCCGTCGCGCGCCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 901
 Db 346 CTACTCTCGTCCGCGCGCCACTTAACCTACCTTCCGTCGCGCGCGCACTTGCCTAC 287
 QY 902 CTCGCTCCGCGCGCGCGCCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961
 Db 286 CTCGCTCCGCGCGCGCGCCACTTAACCTACCTTCCGTCGCGCGCGCACTTGCCT 227
 QY 962 TGCCCCCGCGCAGAGAGCGGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1021
 Db 226 TACCCCGCGCGCAGAAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 167
 QY 1022 ACCTGGGCGAGTCAACACATCTTCCGCTGCGCGCTGAGAGAGACGCGCGCGCA 1081
 Db 166 ACCTAAACAAATCCAGACATCTTCCGCTGAGAGCTCAAAATAAAGCGCGCA 107
 QY 1082 CCGCAGCCTGCTCGGTATTTCGGCATTCAGGCGCGCTACGCTTGAGAGTGACGCGCG 1141
 Db 106 CGACACCTTATCGATTTCGACATACAAACGCGCTACACCTTAAAAAATAACGCGCA 47
 QY 1142 GCGCTGCGGAGCGGTGCGGAGAGGAGGCGCGGT 1182
 Db 46 ACCCTACGAAACGAATAACGAAAAAAGGAAAAACGCGACT 6

RESULT 9
 ABK40089
 ID ABK40089 standard; DNA; 1186 BP.
 XX
 AC ABK40089;
 XX
 DT 21-MAY-2002 (first entry)
 XX Human chemically pretreated gene sequence #86 strand 1.
 DE
 XX Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;
 KW cytosine; ALDH6; CYP11A; CYP11B; CYP11C; DPYD; EPHX2; OCLN; TXNRD1;
 KW UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.
 XX Homo sapiens.
 OS
 XX WO200202806-A2.
 PN
 XX 10-JAN-2002.
 PD

XX 29-JUN-2001; 2001WO-EP07470.
 PF
 XX 30-JUN-2000; 2000DE-1032529.
 PR
 PR 01-SEP-2000; 2000DE-1043826.
 XX (EPIG-) EPIGENOMICS AG.
 PA
 PI Olek A, Piepenbrock C, Berlin K;
 PI WPI, 2002-154757/20.
 DR
 XX New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
 PT useful for detecting cytosine methylation state of genes associated
 with pharmacogenomics and for therapy of diseases e.g. cancer
 XX
 XX Claim 1; SEQ ID No 171; 24pp; English.
 XX
 XX The invention relates to a nucleic acid comprising a sequence at
 CC least 18 bases in length of a segment of the chemically pretreated DNA
 CC of genes associated with pharmacogenomics according to one of the
 CC sequences of the genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B
 CC (NM_000497), CYP3A3 (NM_000776 and NM_017460), DPYD (NM_000110), EPHX2
 CC (NM_001979), OCLN (NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360),
 CC MRP (NM_004996, NM_019900, NM_019901, NM_019902, NM_019862, NM_019898,
 CC NM_019899) and their complementary sequences, or a sequence (S1) chosen
 CC from 87 sequences and their complements. The chemical pretreatment
 CC is bisulphite treatment to convert cytosines (but not methyl-cytosines)
 CC into uracils. Also included are an oligomer (II) in particular an
 CC oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in
 CC each case at least one base sequence having a length of 9 nucleotides
 CC which hybridises to or is identical to a chemically pretreated DNA of
 CC genes associated with pharmacogenomics and their complements, arranged in
 CC an array for analysing diseases associated with the methylation state
 CC (CpG) and/or detecting SNPs (single nucleotide polymorphisms)
 CC of the 87 sequences. The oligomers may also be used as PCR primers.
 CC The set of 87 nucleic acids and their complements is useful for diagnosis
 CC and therapy of solid tumours and cancer. The present sequence
 CC represents one the 87 DNA sequences or its complement.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 1186 BP; 224 A; 98 C; 347 G; 517 T; 0 other;

Query Match 59.8%; Score 71.0; DB 24; Length 1186;
 Best Local Similarity 75.1%; Pred. No. 1.7e-114;
 Matches 892; Conservative 0; Mismatches 294; Indels 1; Gaps 1;

QY 1 GATCGCGCCACTGCACCTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 60
 Db 1 GATCGCGCTATTGATTTTGTGTTGGTGAGAGCGAGATTTGTTTAAAAA 60
 QY 61 AAAAAAGCCGAGGCTCAAAACAAAAACCTCGGAAAGCCCTGGCGGCTTTTTTT 120
 Db 61 AAAAAAGTCGTAGGTTTAAATAAAAAATTTGCGAAAAGTTTGGCGGTTTTTT 120
 QY 121 TTTTTTTTTTTTTTTTTTTTGGGACAGTCTGTCTGTGCGCCAGGCTGAGTACATG 180
 Db 121 TTTTTTTTTTTTTTTTTTTTGGGATAGTCTTTGTTTGTGTTAGTTAGTATAAT 180
 QY 181 CGGATCTGGCTCACTCGAAACCTCTGCTCCAGGTTCAACCAATCTTCTGCCTCAG 240
 Db 181 CGGATTTGGTTTATCTAAATTTTCTTTTGTAGTTTAAAGTAAATTTTGTAGTT 240
 QY 241 TCCCAAGTAGCCACCGCCCACTAAATTTTGTGTAATTTTGTAGTAGACGCGGGTTT 300
 Db 241 TTTTAAGTAGTTATTACGTTTAGTTAAATTTTGTATTTTGTAGTAGACGCGGGTTT 300
 QY 301 CATGTTGTCCAGGCTGTCTNGAACTCTCAGCTCAGCTGATCCACCCGCTCGGCC 360
 Db 301 TATGTTGTTTAGTTGTTTGTGTAATTTTGTAGTTTGTAGTTTATTCGTTTCGTT 360

b 1 TTTAGAGCGGGGCTCTGCGATTCCTCGTAGTAGAGAGCGGAGCGCGAGCGTGC 60
 Y 797 CCGGAAAAGCGCGCGAAGGGTCTCTGCGACCGCGCCACTTGGCGCTGCTCCGTCGCCG 856
 b 61 CCGGAAAAGCGCGCGAAGGGTCTCTGCGATTCCTCGTAGTAGAGAGCGGAGCGTGC 92
 Y 857 CCGGCACTTGGCTGCTCTGCGACCGCGCGCCACTTGGCGCTGCTCCGTCGCCGCGCC 916
 b 93 CCGGCTATTGCTGCTCTGCGACCGCGCGCCACTTGGCGCTGCTCCGTCGCCGCGCC 152
 Y 917 GCGCGCGCATCGCTGCTGCGCGCTGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCG 976
 b 153 GTCGCGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 212
 Y 977 AGGAGCGGAGCG 1036
 b 213 AGGAGCGGAGCG 272
 Y 1037 AACACATCTCTGCTGCGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1096
 b 273 AATAATATTTTCTGCTGCGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 332
 Y 1097 TATTGCGCATCGCGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1156
 b 333 TATTGCGCATCGCGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 392
 Y 1157 GTGCGCGGAGCG 1187
 b 393 GTGCGCGGAGCG 423

RESULT 13
 ABQ39527/c
 ID ABQ39527 standard; DNA; 764 BP.
 XX AC ABQ39527;
 XX DT 12-JUL-2002 (first entry)
 XX OS Oligonucleotide for detecting cytosine methylation SEQ ID NO 26118.
 XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 XX KW drug; side effect; cancer; central nervous system; cardiovascular;
 XX KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 XX KW SNP; cell differentiation; ds.
 XX OS Homo sapiens.
 XX WO200218632-A2.
 XX PD 07-MAR-2002.
 XX XX 01-SEP-2001; 2001WO-EP10074.
 XX XX 01-SEP-2000; 2000DE-1043826.
 XX XX 05-SEP-2000; 2000DE-1044543.
 XX XX (EPIC-) EPIGNOMICS AG.
 XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;
 XX WI; 2002-371829/40.
 XX XX Determining the degree of cytosine methylation in genomic DNA, useful
 XX XX for diagnosis and prognosis, comprises selective hybridization of
 XX XX amplicons from chemically treated DNA -
 XX XX Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX XX This invention describes a novel method for determining the degree of
 XX XX methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 XX XX genomic sample of DNA. The sample is treated chemically to convert

CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP/s); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABQ3410-ABQ54121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.

XX Sequence 764 BP; 274 A; 283 C; 98 G; 109 T; 0 other;

Query Match 19.1%; Score 226.6; DB 24; Length 764;
 Best Local Similarity 71.8%; Pred. No. 1e-30;
 Matches 324; Conservative 0; Mismatches 99; Indels 28; Gaps 1;

QY 737 TCTAGAGCGGGGCG 796
 Db 764 TTTAGAGCGGGGCTCTGCGATTCCTCGTAGTAGAGAGCGGAGCGGAGCGGAG 705
 QY 797 CCGGAAAAGCGCGCGGAAAGGGGCTCTGCCACCGCGCGCGCGCGCGCGCGCG 856
 Db 704 CCGGAAAAGCGCGCGGAAAGGGGCTCTGCCACCGCGCGCGCGCGCGCGCGCG 873
 QY 857 CGCGCGCATCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 916
 Db 672 CGCGTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
 QY 917 GCGCGCGCATCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 976
 Db 612 CTCGCGTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 553
 QY 977 AGGAGCGGAGCG 1036
 Db 552 AGGAGCGGAGCG 493
 QY 1037 AACACATCTCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 1096
 Db 492 AATATATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 433
 QY 1097 TATTGCGCATCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 1156
 Db 432 TATTGCGCATCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 373
 QY 1157 GTGCGCGGAGCG 1187
 Db 372 GTGCGCGGAGCG 342

RESULT 14
 ABQ39528/c
 ID ABQ39528 standard; DNA; 764 BP.
 XX AC ABQ39528;
 XX DT 12-JUL-2002 (first entry)
 XX XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 26119.
 XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 XX KW drug; side effect; cancer; central nervous system; cardiovascular;
 XX KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 XX KW SNP; cell differentiation; ds.
 XX OS Homo sapiens.

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

M nucleic - nucleic search, using sw model

un on: November 28, 2002, 18:20:44 ; Search time 3018.72 Seconds
(without alignments)
11443.601 Million cell updates/sec

itle: US-09-963-333-7

effect score: 1187

sequence: 1 gatcgccactgcactcca.....gaggagcgccgtgggga 1187

coring table: IDENTIFY NUC

Gapop 10.0 , Gapext 1.0

sarched: 2054640 seqs, 14551402878 residues

otal number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1167	98.3	1186	9	HUMTS	D00517 Homo sapien
2	1167	98.3	18596	6	AR144965	AR144965 Sequence
3	1167	98.3	18596	6	AX050451	AX050451 Sequence
4	1167	98.3	18596	6	AX330682	AX330682 Sequence
5	1167	98.3	18596	6	AX330906	AX330906 Sequence
6	1167	98.3	18596	6	AX335755	AX335755 Sequence
7	1167	98.3	18596	6	AX397714	AX397714 Sequence
8	1167	98.3	18596	6	AX408943	AX408943 Sequence
9	1167	98.3	18596	9	HUMTS1	D00596 Homo sapien
10	1149.8	96.9	152711	9	AP001178	AP001178 Homo sapi
11	1006	84.8	165742	2	AC020697	AC020697 Homo sapi
12	770.6	64.9	1186	6	AX348476	AX348476 Sequence
13	710	59.8	1186	6	AX348476	AX348476 Sequence
14	299.4	25.2	1536	6	AR103889	AR103889 Sequence
15	299.4	25.2	1536	6	AX397715	AX397715 Sequence
16	299.4	25.2	1536	6	AX482020	AX482020 Sequence
17	299.4	25.2	1536	6	HSTSYN1	X02308 Human mRNA
18	288.8	24.3	3298	6	AX281791	AX281791 Sequence
19	278.6	23.5	1567	9	AB062290	AB062290 Homo sapi
20	278.6	23.5	1569	9	BC013919	BC013919 Homo sapi
21	275.8	23.2	1533	9	BC002567	BC002567 Homo sapi
22	228.2	19.2	233454	9	AC005703	AC005703 Homo sapi
23	225.2	19.0	80565	9	AL357374	AL357374 Human DNA
24	224.6	18.9	193487	2	AC127455	AC127455 Homo sapi
25	224.4	18.9	190871	9	AC044797	AC044797 Homo sapi
26	222.6	18.8	202827	9	AC016907	AC016907 Homo sapi
27	221.8	18.7	193487	2	AC127455	AC127455 Homo sapi
28	221.6	18.7	160536	9	AL355984	AL355984 Human DNA
29	221.2	18.6	179641	9	AC098591	AC098591 Homo sapi
30	221	18.6	150296	9	HUU91326	U91326 Homo sapien
31	220.6	18.6	97137	9	AC004745	AC004745 Homo sapi
32	220.4	18.6	132948	9	HS349A12	AL033520 Human DNA
33	219.6	18.5	146407	9	AC004953	AC004953 Homo sapi
34	219.6	18.5	195773	2	AC073134	AC073134 Homo sapi
35	219.4	18.5	104289	2	HS104D14	AL035554 Homo sapi
36	219	18.4	43497	9	HS695020A	AL049856 Human DNA
37	219	18.4	211176	2	HSDA109B7	AL121672 Homo sapi
38	218.8	18.4	156602	9	AC012487	AC012487 Homo sapi
39	218.6	18.4	148954	9	AL590640	AL590640 Human DNA
40	218.2	18.4	180346	9	AC092615	AC092615 Homo sapi
41	218	18.4	142667	9	AC125387	AC125387 Homo sapi
42	218	18.4	181610	9	AC100860	AC100860 Homo sapi
43	217.8	18.3	168006	2	AC126227	AC126227 Papio cyn
44	217.6	18.3	305000	9	HSXDPA	AL590762 Homo sapi
45	217.4	18.3	63797	2	AC127453	AC127453 Homo sapi

ALIGNMENTS

RESULT 1

HUMTS

LOCUS HUMTS

DEFINITION Homo sapiens gene for thymidylate synthase, exon 1, partial cds.

ACCESSION D00517

VERSION D00517.1

KEYWORDS GI:220133

SOURCE Homo sapiens DNA, clone:lambda HTS-1.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 1186)

Takeishi,K., Kaneda,S., Ayusawa,D., Shimizu,K., Gotoh,O. and

Seno,I.

TITLE Human thymidylate synthase gene: isolation of phage clones which

Pred. No. is the number of results predicted by chance to have a

cover a functionally active gene and structural analysis of the region upstream from the translation initiation codon
J. Biochem. 106 (4), 575-583 (1989)

90110051
These data kindly submitted in computer readable form by: Keiichi Takeishi
University of Shizuoka School of Food and Nutritional Sciences 395
Yada
Shizuoka-shi,
Shizuoka-ken 422
Japan
Phone: 0542-64-5540
Fax: 0542-64-5099.

FEATURES
source
1. .1186
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/clone="lambda HTS-1"
/note="Sau3AI site-from a genomic library of the mouse cell transformant CO"
115..400
repeat_unit
/note="Alu sequence"

prim_transcript
751..>1186
/notes="thymidylate synthase mRNA and intron"
exon
751..1128
/number=1
repeat_unit
785..812
/notes="inverted sequence to TTR"
repeat_unit
827..916
/notes="triple tandem repeat (TTR)"
924..>1128
/EC_number="2.1.1.45"
/note="exon1 part"

CDS
/product="thymidylate synthase"
/protein_id="BAA00404.1"
/db_xref="GI:2160415"
/translation="MPVAGSLPRPLPPAQRDAEPHPHGLQYLQIQLHRCG
VRKDRGTGTLVFMQARYSLR"
1129..>1186
/number=1

intron
224 a 386 c 347 g 229 t
BASE COUNT
ORIGIN
Query Match 98.3%; Score 1167; DB 9; Length 1186;
Best Local Similarity 99.2%; Pred. No. 4.6e-201;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Qy 1 GATCGCGCACTGCACCTCCAGCTGGGTGAGAGAGCGAGACTCTGTCTCAAAAAA 60
Db 1 GATCGCGCACTGCACCTCCAGCTGGGTGAGAGAGCGAGACTCTGTCTCAAAAAA 60
Qy 61 AAAAAGACCGCAGGGCTCAACAAACAAACCTCGGAAGCCCTGGCGGTCTTTTTTT 120
Db 61 AAAAAGACCGCAGGGCTCAACAAACAAACCTCGGAAGCCCTGGCGGTCTTTTTTT 120
Qy 121 TTTTCTTTTTTTTTTTTGGGACAGTCTGTCTGTGCCAGGCTGGAGTACAATGGT 180
Db 121 TTTTCTTTTTTTTTTTTGGGACAGTCTGTCTGTGCCAGGCTGGAGTACAATGGT 180
Qy 181 CGGATCTGGCTCACTCAACCTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 240
Db 181 CGGATCTGGCTCACTCAACCTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 240
Qy 241 TCCAGTAGCCACACCGCCAGTAAATTTTGTANTTTTGTAGAGACGGGGTTTAC 300
Db 241 TCCAGTAGCCACACCGCCAGTAAATTTTGTANTTTTGTAGAGACGGGGTTTAC 300
Qy 301 CATGTTGTCAGGCTGGTCTNGAACTCTGACCTCAGGTGATCCACCGCTCGGCCCC 360
Db 301 CATGTTGTCAGGCTGGTCTNGAACTCTGACCTCAGGTGATCCACCGCTCGGCCCC 360

Qy 361 CAAAGTACTAGATTACAGCGGTGAGCCACCGCGTCCAGCGCCCTGGCGGTTTTTAATCA 420
Db 361 CAAAGTACTAGATTACAGCGGTGAGCCACCGCGTCCAGCGCCCTGGCGGTTTTTAATCA 420
Qy 421 AGTAGAAAAGCTGCAATATACCACTTGTCTTCNGTTGNCNTTCAGTGAGAACAGAAATGG 480
Db 421 AGTAGAAAAGCTGCAATATACCACTTGTCTTCNGTTGNCNTTCAGTGAGAACAGAAATGG 479
Qy 481 AAATGCAATNCNTTATTAGTTGTAGAAAACAGATCTCAACACAGCAGTTTGTGNGACAAG 540
Db 480 AAATGCAATNCNTTATTAGTTGTAGAAAACAGATCTCAACACAGCAGTTTGTGNGACAAG 539
Qy 541 ACCGAGGAAAAACGTGGGAACTGTGTCTGTGCTTTAGAGAGGCGCGGTCCAGCAGCGG 600
Db 540 ACCGAGGAAAAACGTGGGAACTGTGTCTGTGCTTTAGAGAGGCGCGGTCCAGCAGCGG 599
Qy 601 TTCCCAAGAGGCGCAGTCTCTTCCNGCCACCGACCTGCNTCCAGGTTCCCGGTTTCT 660
Db 600 TTCCCAAGAGGCGCAGTCTCTTCCNGCCACCGACCTGCNTCCAGGTTCCCGGTTTCT 659
Qy 661 AAGACTCTCAGCTGTGGCCCTGGGCTCCGTTCTGTGCCACACCCCGTCTCTGCGTTTC 720
Db 660 AAGACTCTCAGCTGTGGCCCTGGGCTCCGTTCTGTGCCACACCCCGTCTCTGCGTTTC 719
Qy 721 CCCTCGCGCAGCTCTCTAGAGCGGGGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Db 720 CCCTCGCGCAGCTCTCTAGAGCGGGGCGCGCGGAGGAGGAGGAGGAGGAGGAGG 779
Qy 781 GAGCGCGGAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
Db 780 GAGCGCGGAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 839
Qy 841 CTGCGCTCTCGCTCCGCGCGCGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 840 CTGCGCTCTCGCTCCGCGCGCGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 899
Qy 901 CTGCGCTCTCGCTCCGCGCGCGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 900 CTGCGCTCTCGCTCCGCGCGCGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 959
Qy 961 TTGCGCGCGCGCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
Db 960 TTGCGCGCGCGCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1019
Qy 1021 TACTGCGCGCAGATCCAAACATCTCTCGCTGCGGCTCAGAAAGGAGCAGCAGCAGCGGC 1080
Db 1020 TACTGCGCGCAGATCCAAACATCTCTCGCTGCGGCTCAGAAAGGAGCAGCAGCAGCGGC 1079
Qy 1081 ACCGCGACCTCTCGTATTCTGCGATGCGAGCGCGCTACAGCTTGAAGGTGACCGCGCG 1140
Db 1080 ACCGCGACCTCTCGTATTCTGCGATGCGAGCGCGCTACAGCTTGAAGGTGACCGCGCG 1139
Qy 1141 GGCGCTGCGGAGCGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1187
Db 1140 GGCGCTGCGGAGCGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1186

RESULT 2
AR144965
LOCUS AR144965 18596 bp DNA linear PAT 09-AUG-2001
DEFINITION Sequence 11 from patent US 6210950.
ACCESSION AR144965
VERSION AR144965.1 GI:15106832
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 18596)
AUTHORS Johnson, W.G. and Stenroos, B.Scott.
TITLE Methods for diagnosing, preventing, and treating developmental disorders due to a combination of genetic and environmental factors
JOURNAL Patent: US 6210950-A 11 03-APR-2001;
FEATURES Location/Qualifiers

[illegible]


```
source 1..18596
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4521 a 3991 c 4479 g 5605 t
ORIGIN
Query Match 98.3%; Score 1167; DB 6; Length 18596;
Best Local Similarity 99.2%; Pred. No. 3.8e-201;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 1 GATCGCGCCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 60
Db 78 GATCGCGCCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 137
QY 61 AAAAAAGACCGCGGCTCAAAACAAAAAAGCTCGGAAAGCCCTGGCGGTCTTTT 120
Db 138 AAAAAAGACCGCGGCTCAAAACAAAAAAGCTCGGAAAGCCCTGGCGGTCTTTT 197
QY 121 TTTTCTTTTCTTTTCTTTTGGGACAGTCTTGTCTGTGCGCCAGGCTGAGTACATG 180
Db 121 TTTTCTTTTCTTTTCTTTTGGGACAGTCTTGTCTGTGCGCCAGGCTGAGTACATG 257
QY 181 CGGATCTTGGCTCACTCAAAACCTCTGCTCCAGGTTCCAGGTTCCAGGTTCCAG 240
Db 181 CGGATCTTGGCTCACTCAAAACCTCTGCTCCAGGTTCCAGGTTCCAGGTTCCAG 317
QY 241 TCCCAAGTAGCCACCGAGCTGATTTTGTANTTTTAGTAGAGAGCGGGTTTCAC 300
Db 241 TCCCAAGTAGCCACCGAGCTGATTTTGTANTTTTAGTAGAGAGCGGGTTTCAC 377
QY 301 CATGTGTCCAGGCTGTCTGACCTCAGTGTGATCCACCGGCTCGGCGCC 360
Db 301 CATGTGTCCAGGCTGTCTGACCTCAGTGTGATCCACCGGCTCGGCGCC 437
QY 361 CAAAGTACTAGGATTTACAGGCGTGAAGCCACCGCTCCAGCGCTCGGCGTTTAA 420
Db 361 CAAAGTACTAGGATTTACAGGCGTGAAGCCACCGCTCCAGCGCTCGGCGTTTAA 497
QY 421 AGTAAAAAGCTGCAATATACCACTGTCTTGGTGGTGGTGGTGGTGGTGGTGG 480
Db 421 AGTAAAAAGCTGCAATATACCACTGTCTTGGTGGTGGTGGTGGTGGTGGTGG 556
QY 481 AATGCAATATCCTTTTATAGTTAGTGAACAGATCTCAAAACAGGTTTGTGACAG 540
Db 481 AATGCAATATCCTTTTATAGTTAGTGAACAGATCTCAAAACAGGTTTGTGACAG 616
QY 541 ACCGAGGAAACGTTGGGAACTGTCTGCTGGCTTAGAGAAAGCGCGGTCCAGACG 600
Db 541 ACCGAGGAAACGTTGGGAACTGTCTGCTGGCTTAGAGAAAGCGCGGTCCAGACG 676
QY 601 TTCCCAAGGCGGAGTCTTCCCGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 601 TTCCCAAGGCGGAGTCTTCCCGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
QY 661 AAGACTCTCAGTGTGGGCTCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 AAGACTCTCAGTGTGGGCTCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
QY 721 CCGCTGCGGACGCTCTTAGAGCGGGGCGCGGAGGCTGCTGCTGCTGCTGCTGCTG 780
Db 721 CCGCTGCGGACGCTCTTAGAGCGGGGCGCGGAGGCTGCTGCTGCTGCTGCTGCTG 856
QY 781 GAGCGGGAGGCGCGCGGAAAGGCGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 781 GAGCGGGAGGCGCGCGGAAAGGCGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTG 916
QY 841 CTTGCTTCCGTCGCGCGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 841 CTTGCTTCCGTCGCGCGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
QY 901 CCTCGGTCCTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 CCTCGGTCCTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1036
```

```
QY 961 TTGCCCCCGCGCCACAGGAGCGGAGCGCGAGCGCGCTCCGCGCACGCGGAGCTGCAG 1020
Db 1037 TTGCCCCCGCGCCACAGGAGCGGAGCGCGAGCGCGCTCCGCGCACGCGGAGCTGCAG 1096
QY 1021 TACCTGGGCGAGATCCAAACATCTTCCGCTGCGGCGCTCAGGAAGGACGACCGCGCG 1080
Db 1097 TACCTGGGCGAGATCCAAACATCTTCCGCTGCGGCGCTCAGGAAGGACGACCGCGCG 1156
QY 1081 ACGGCAACCTCTGTGCTGATTCGCGATGACAGCGCGCTACAGCTGAGAGGTGACCGCG 1140
Db 1157 ACGGCAACCTCTGTGCTGATTCGCGATGACAGCGCGCTACAGCTGAGAGGTGACCGCG 1216
QY 1141 GCGCCCTCGGAGCGGCTGCGGAAAGGAGGAGCGCGCGGTGGGGA 1187
Db 1217 GCGCCCTCGGAGCGGCTGCGGAAAGGAGGAGCGCGCGGTGGGGA 1263
RESULT 7
AX397714 18596 bp DNA linear PAT 18-MAY-2002
LOCUS
DEFINITION Sequence 140 from Patent WO210366.
ACCESSION AX397714
VERSION AX397714.1 GI:21068436
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Chen, H., Freilmer, N. B. and Novak, T.
TITLE Methods and compositions for diagnosing and treating chromosome-18P
related disorders
JOURNAL Patent: WO 0210366-A 140 07-FEB-2002;
Millennium Pharmaceuticals, Inc. (US); The Regents of The
University of California (US)
FEATURES
source 1..18596
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4521 a 3991 c 4479 g 5605 t
ORIGIN
Query Match 98.3%; Score 1167; DB 6; Length 18596;
Best Local Similarity 99.2%; Pred. No. 3.8e-201;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 1 GATCGCGCCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 60
Db 78 GATCGCGCCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 137
QY 61 AAAAAAGCGCGAGGCTCAAAACAAAAAAGCTCGGAAAGCCCTGGCGGTCTTTT 120
Db 138 AAAAAAGCGCGAGGCTCAAAACAAAAAAGCTCGGAAAGCCCTGGCGGTCTTTT 197
QY 121 TTTTCTTTTCTTTTCTTTTGGGACAGTCTTGTCTGTGCGCCAGGCTGAGTACATG 180
Db 121 TTTTCTTTTCTTTTCTTTTGGGACAGTCTTGTCTGTGCGCCAGGCTGAGTACATG 257
QY 181 CGGATCTTGGCTCACTCAAAACCTCTGCTCCAGGTTCCAGGTTCCAGGTTCCAG 240
Db 181 CGGATCTTGGCTCACTCAAAACCTCTGCTCCAGGTTCCAGGTTCCAGGTTCCAG 317
QY 241 TCCCAAGTAGCCACCGAGCTGATTTTGTANTTTTAGTAGAGAGCGGGTTTCAC 300
Db 241 TCCCAAGTAGCCACCGAGCTGATTTTGTANTTTTAGTAGAGAGCGGGTTTCAC 377
QY 301 CATGTGTCCAGGCTGTCTGACCTCAGTGTGATCCACCGGCTCGGCGCC 360
Db 301 CATGTGTCCAGGCTGTCTGACCTCAGTGTGATCCACCGGCTCGGCGCC 437
QY 361 CAAAGTACTAGGATTTACAGGCGTGAAGCCACCGCTCCAGCGCTCGGCGTTTAA 420
Db 361 CAAAGTACTAGGATTTACAGGCGTGAAGCCACCGCTCCAGCGCTCGGCGTTTAA 497
QY 421 AGTAAAAAGCTGCAATATACCACTGTCTTGGTGGTGGTGGTGGTGGTGGTGG 480
Db 421 AGTAAAAAGCTGCAATATACCACTGTCTTGGTGGTGGTGGTGGTGGTGGTGG 556
QY 481 AATGCAATATCCTTTTATAGTTAGTGAACAGATCTCAAAACAGGTTTGTGACAG 540
Db 481 AATGCAATATCCTTTTATAGTTAGTGAACAGATCTCAAAACAGGTTTGTGACAG 616
QY 541 ACCGAGGAAACGTTGGGAACTGTCTGCTGGCTTAGAGAAAGCGCGGTCCAGACG 600
Db 541 ACCGAGGAAACGTTGGGAACTGTCTGCTGGCTTAGAGAAAGCGCGGTCCAGACG 676
QY 601 TTCCCAAGGCGGAGTCTTCCCGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 601 TTCCCAAGGCGGAGTCTTCCCGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
QY 661 AAGACTCTCAGTGTGGGCTCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 AAGACTCTCAGTGTGGGCTCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
QY 721 CCGCTGCGGACGCTCTTAGAGCGGGGCGCGGAGGCTGCTGCTGCTGCTGCTGCTG 780
Db 721 CCGCTGCGGACGCTCTTAGAGCGGGGCGCGGAGGCTGCTGCTGCTGCTGCTGCTG 856
QY 781 GAGCGGGAGGCGCGCGGAAAGGCGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 781 GAGCGGGAGGCGCGCGGAAAGGCGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTG 916
QY 841 CTTGCTTCCGTCGCGCGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 841 CTTGCTTCCGTCGCGCGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
QY 901 CCTCGGTCCTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 CCTCGGTCCTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1036
```

438 CAAAGTACTAGGANTACAGCGCTGAGCCACCGCGTCCAGCGCCCTGCGGCTTTTAATCA 497
421 AGTAAAAAGCTGCAATATACACATGCTTCNGTGTGCTGNTTCACTGAGAACGAAAGATGG 480
498 AGTAAAAAGCTGCAATATACACATGCTTCGCTGTC-TTCAGTGAGAACGAAAGATGG 556
481 AAATGCAATCNCCTTATATAGTTGTAGGAAACAGATCTCAAAACAGAGTTTGTGACAG 540
557 AAATGCAATCNCCTTATATAGTTGTAGGAAACAGATCTCAAAACAGAGTTTGTGACAG 616
541 ACCCGAGGAAACGTCGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
617 ACCCGAGGAAACGTCGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676
601 TTCCCAAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
677 TTCCCAAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
661 AAGACTCTAGCTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
737 AAGACTCTAGCTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 796
721 CCCCTGCGCAGCTCTCTAGAGCGGGGCGCGCGAGCCCGCGAGCCCGCGAGGAGCGG 780
797 CCCCTGCGCAGCTCTCTAGAGCGGGGCGCGCGAGCCCGCGAGCCCGCGAGGAGCGG 856
781 GAGCGCGGAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
857 GAGCGCGGAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 916
841 CTTGCTTCCCG 900
917 CTTGCTTCCCG 976
901 CTTGCTTCCCG 960
977 CTTGCTTCCCG 1036
961 TTGCG 1020
1037 TTGCG 1096
1021 TACCTGGGCGAGATCAACACATCTCCGCTGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGG 1080
1097 TACCTGGGCGAGATCAACACATCTCCGCTGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGG 1156
1081 ACCGCGACCTGTCGCTATTCGCGATGACGCGCGCTACAGCTGAGAGGTGACGCGCG 1140
1157 ACCGCGACCTGTCGCTATTCGCGATGACGCGCGCTACAGCTGAGAGGTGACGCGCG 1216
1141 GCGCTTCCGCGAGCGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1187
1217 GCGCTTCCGCGAGCGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1263

RESULT 8
AX408943 LOCUS AX408943 18596 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 1590 from Patent WO0229103.
ACCESSION AX408943
VERSION AX408943.1 GI:21441648
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Alvares, C., Horns, D., Peres-da-Silva, S. and Vockley, J.G.
TITLE Gene expression profiles in liver cancer
PATENT: WO 0229103-A 1590 11-APR-2002;
JOURNAL GENE LOGIC INC (US)
FEATURES
LOCATION/Qualifiers
source 1..18596

/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. D00596"
BASE COUNT 4521 a 3991 c 4479 g 5605 t
ORIGIN
Query Match 98.3%; Score 1167; DB 6; Length 18596;
Best Local Similarity 99.2%; Pred. No. 3.8e-201; Mismatches 8; Indels 1; Gaps 1;
Matches 1178; Conservative 0;
QY 1 GATCGCGCACTGCACTCCAGCCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 60
DB 78 GATCGCGCACTGCACTCCAGCCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 137
QY 61 AAAAAAGCCGCGAGGCTCAAAACAAAAAAGCTGGAAGCCCTGGCGGCTTTT 120
DB 138 AAAAAAGCCGCGAGGCTCAAAACAAAAAAGCTGGAAGCCCTGGCGGCTTTT 197
QY 121 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 180
DB 198 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 257
QY 181 CGGATCTTGGCTCACTGCACTCCAGCCTGGGTGAGAGCGAGACTCTGTCTCA 240
DB 258 CGGATCTTGGCTCACTGCACTCCAGCCTGGGTGAGAGCGAGACTCTGTCTCA 317
QY 241 TCCCAAGTAGCAGCAGCCAGCTAAATTTTGTATTTTGTATTTTGTATAGAGA 300
DB 318 TCCCAAGTAGCAGCAGCCAGCTAAATTTTGTATTTTGTATTTTGTATAGAGA 377
QY 301 CATGTTGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 378 CATGTTGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 437
QY 361 CAAAGTACTAGATTTACAGGCGTGAGCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 438 CAAAGTACTAGATTTACAGGCGTGAGCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 497
QY 421 AGTAGAAAAGCTGCAATTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 498 AGTAGAAAAGCTGCAATTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 556
QY 481 AAATGCAATCNCCTTATATAGTTGTAGGAAACAGATCTCAAAACAGAGTTTGTGACAG 540
DB 557 AAATGCAATCNCCTTATATAGTTGTAGGAAACAGATCTCAAAACAGAGTTTGTGACAG 616
QY 541 ACCCGAGGAAACGTCGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 617 ACCCGAGGAAACGTCGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676
QY 601 TTCCCAAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 677 TTCCCAAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 736
QY 661 AAGACTCTAGCTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 737 AAGACTCTAGCTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 796
QY 721 CCCCTGCGCAGCTCTCTAGAGCGGGGCGCGCGAGCCCGCGAGCCCGCGAGGAGCGG 780
DB 797 CCCCTGCGCAGCTCTCTAGAGCGGGGCGCGCGAGCCCGCGAGCCCGCGAGGAGCGG 856
QY 781 GAGCGCGGAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
DB 857 GAGCGCGGAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 916
QY 841 CTTGCTTCCCG 900
DB 917 CTTGCTTCCCG 976
QY 901 CTTGCTTCCCG 960
DB 977 CTTGCTTCCCG 1036

QY 961 TTGCCCCCGCGGACAGGAGCGGACGCGGAGCGCGGTCCGCCACGCGGAGCTGCAG 1020
 Db 1037 TTGCCCCCGCGGACAGGAGCGGACGCGGAGCGCGGTCCGCCACGCGGAGCTGCAG 1096
 QY 1021 TACTCTGGGCGCATCACAACATCTTCGCTGCGGCGGTCTAGGAAGACGACCGACGCGGC 1080
 Db 1097 TACTCTGGGCGCATCACAACATCTTCGCTGCGGCGGTCTAGGAAGACGACCGACGCGGC 1156
 QY 1081 ACCGCGACCTGTGCGTATTCGCGATGCAAGCGCGCTACAGCTGAGAGTGCACCGCGC 1140
 Db 1157 ACCGCGACCTGTGCGTATTCGCGATGCAAGCGCGCTACAGCTGAGAGTGCACCGCGC 1216
 QY 1141 GCGCCCTCGGAGCGGTGCGCGGAGGAGGCGCGGCTGGGA 1187
 Db 1217 GCGCCCTCGGAGCGGTGCGCGGAGGAGGCGCGGCTGGGA 1263

RESULT 9

HUMTS1 18596 bp DNA linear PRI 14-APR-2000
 LOCUS Homo sapiens gene for thymidylate synthase, exons 1, 2, 3, 4, 5, 6,
 DEFINITION 7, complete cds.

ACCESSION D00596

VERSION D00596.1 GI:220135

KEYWORDS thymidylate synthase.

SOURCE Homo sapiens DNA, clone:lambdaHTS-1 and lambdaHTS-3.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Kaneda, S., Nalbantoglu, J., Takeishi, K., Shimizu, K., Gotoh, O.,

Seno, T., and Aoyama, D.

Structural and functional analysis of the human thymidylate

synthase gene

J. Biol. Chem. 265 (33), 20277-20284 (1990)

MEDLINE

91056070

COMMENT

These data kindly submitted in computer readable form by: Sumiko

Kaneda

National Institute of Genetics

1111 Yata

Mishima 411

Japan

Phone: +81-559-72-2732

Fax: +81-559-71-3651.

Location/Qualifiers

1. 18596

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="18"

/map="18p11.32"

/clones="lambdaHTS-1 and lambdaHTS-3"

repeat_unit 1. .148

/note="Alu sequence"

repeat_unit 202. .477

/note="Alu sequence"

prim_transcript 822. .16246

/note="thymidylate synthase mRNA and introns"

prim_transcript 824. .16246

/note="thymidylate synthase mRNA and introns"

prim_transcript 828. .16246

/note="thymidylate synthase mRNA and introns"

prim_transcript 841. .16246

/note="thymidylate synthase mRNA and introns"

repeat_unit 862. .889

/note="inverted repeat"

repeat_unit 904. .993

/note="triple tandem repeats"

join(1001..1205,2895..2968,5396..5570,11843..11944,

13449..13624,14133..14204,15613..15750)

/EC_number="2.1.1.45"

/codon_start=1

/product="thymidylate synthase"

/protein_id="BAA00472.1"
 /db_xref="GI:220136"
 /translation="MPVAGSELPRRLPPAAQERDAERPPHGELOYLQIHLRCG
 VRKDRGTGTLTVFGMQARYSLRDFPLLTTRVFWKGVLELLNFIKGSTNAKELS
 SKGKTIWDANGSRDFLDSLGSTREGLGPVYFQWRHFGAERYDMESYSGQVDO
 LQVDTITKTPDDERLIIMCWNPRDLPLMALPPCHALCOFVNVNSLSQQLQREPRPF
 MGUGVFNTASVALLTYMIAHTGLKPGDFITLGDHAHYILNHIEPLKIQLQREPRPF
 PKLRIILKVEKIDDFKAFDFQIEGYNPHPTIKEMAW"
 <1001..1205

/number=1

1206. .2894

/number=1

1660. .1665

/bound_moiety="Spl"

1890. .1895

/bound_moiety="Spl"

1895. .1900

/bound_moiety="Spl"

2895. .2968

/number=2

2969. .5395

/number=2

3049. .3347

/note="Alu sequence"

4953. .5245

/note="Alu sequence"

5396. .5570

/number=3

5571. .11842

/number=3

5768. .9405

/note="Alu sequence"

10125. .10389

/note="reiterated motif of GGA/TCAT"

11843. .11944

/number=4

11945. .13448

/number=4

12135. .12397

/note="Alu sequence"

12567. .12743

/note="Alu sequence"

12807. .13084

/note="Alu sequence"

13449. .13624

/number=5

13625. .14132

/number=5

13875. .14010

/note="Alu sequence"

14133. .14204

/number=6

14205. .15612

/number=6

14531. .14832

/note="Alu sequence"

15613. .15750

/number=7

15934. .15939

polyA signal

16228. .16233

polyA signal

17262. .17553

repeat_unit

/note="Alu sequence"

BASE COUNT 4521 a 3991 c 4479 g 5605 t

ORIGIN

Query Match 98.3%; Score 1167; DB 9; Length 18596;

Best Local Similarity 99.2%; Pred. No. 3.8e-201;

Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1 GATCGCGCCTGCACCTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAAAA 60

|||||

Db 78 GATCGCGCCTGCACCTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAAAA 137

|||||

61	AAAAGACGGCCAGGGCTCAAA	CAAAAAA	CTCGAAAAAGCCCTGGCGGTTCTTTT	TTTTT	120
138	AAAAGACGGCCAGGGCTCAAA	CAAAAAGACCTCGAAAAAGCCCTGGCGGTTCTTTT	TTTTT	197	
121	TTTTTTTTTTTTTTTTTTTT	TGGGACAGCTCTTGCTCTGTGCCCGACGCTGGAGTACAATGGT		180	
198	TTTTTTTTTTTTTTTTTTTT	TGGGACAGCTCTTGCTCTGTGCCCGACGCTGGAGTACAATGGT		257	
181	CGGATCTGGCTCACTGCAACCT	CTGGCTCCCGACGGTTCAAGCAATTTCTTGCTCGCCAGCC	240		
258	CGGATCTGGCTCACTGCAACCT	CTGGCTCCCGACGGTTCAAGCAATTTCTTGCTCGCCAGCC	317		
241	TCCCAAGTAGCCACACGCC	CCAGCTAATTTTGTANTTTTGTAGTGGAGCGGGGTTTCAAC	300		
318	TCCCAAGTAGCCACACGCC	CCAGCTAATTTTGTANTTTTGTAGTGGAGCGGGGTTTCAAC	377		
301	CATGTTCTCCAGGCTGGTCT	NGAACTCTGACCTCAGGTGAATCCACCGCCCTCGGCCCCC	360		
378	CATGTTCTCCAGGCTGGTCT	TGAACTCTGACCTCAGGTGAATCCACCGCCCTCGGCCCCC	437		
361	CAAAAGTACTAGGATTACAGG	CGGTGAGCAACCGCTCAGCGCCCTGGCGGTTTAAATCA	420		
438	CAAAAGTACTAGGATTACAGG	CGGTGAGCAACCGCTCAGCGCCCTGGCGGTTTAAATCA	497		
421	AGTAGAAAAGCTGCAATTA	ACACTTGTCTGTTGNTTCACTGAGGACCAAGAATGG	480		
498	AGTAGAAAAGCTGCAATTA	ACACTTGTCTGTTGNTTCACTGAGGACCAAGAATGG	556		
481	AAATGCAAAATCNCNTTA	TAGTTGTAGGAAACAGATCTCAACACGACAGTTTGTGACAAAG	540		
557	AAATGCAAAATCNCNTTA	TAGTTGTAGGAAACAGATCTCAACACGACAGTTTGTGACAAAG	616		
541	ACCGCAGGAAAGCTGGGAA	CTGTGCTGTGCTTAGAGAGCGCGGTCGACACAGCG	600		
617	ACCGCAGGAAAGCTGGGAA	CTGTGCTGTGCTTAGAGAGCGCGGTCGACACAGCG	676		
601	TTCCCAAGGGCGCAGTCC	TTCCCGGCCACCGCACTCGTNCFTCAGGTTCCCGGGTTCCT	560		
677	TTCCCAAGGGCGCAGTCC	TTCCCGGCCACCGCACTCGTNCFTCAGGTTCCCGGGTTCCT	736		
661	AAGACTCTCAGCTGTGCGC	CTGGCTCCGTTGTGCCACACCGCTGGTCTCTGGGTTTC	720		
737	AAGACTCTCAGCTGTGCGC	CTGGCTCCGTTGTGCCACACCGCTGGTCTCTGGGTTTC	796		
721	CCCTGGCGGCAACGCTCT	TAGAGCGGGGGCCCGCGACCCCGCCGACGACGAGAGCG	780		
797	CCCTGGCGGCAACGCTCT	TAGAGCGGGGGCCCGCGACCCCGCCGACGACGAGAGCG	856		
781	GAGCGCGGGAACGCGCGG	GAAGAGCGCGGAAGGGGTCTGCGACACCGCGCACTTTGG	840		
857	GAGCGCGGGAACGCGCGG	GAAGAGCGCGGAAGGGGTCTGCGACACCGCGCACTTTGG	916		
841	CTTGCTCTCGTCCGCGG	CCCACTTGTGCTCCGTCGCGCGCGGCGCACTTCGCTCG	900		
917	CTTGCTCTCGTCCGCGG	CCCACTTGTGCTCCGTCGCGCGCGGCGCACTTCGCTCG	976		
901	CCCTCGTCCCGCGCGC	CGCCATGCTGTGCGCGGCTCGGAGCTCGCGCGCGCGCC	960		
977	CCCTCGTCCCGCGCGC	CGCCATGCTGTGCGCGGCTCGGAGCTCGCGCGCGCGCC	1036		
961	TTGCCCCCGCGCGCA	CAGGAGGGGACCGCGAGCGCGTCCGCGCACGAGGAGTGTGAG	1020		
1037	TTGCCCCCGCGCGCA	CAGGAGGGGACCGCGAGCGCGTCCGCGCACGAGGAGTGTGAG	1096		
1021	TACTGGGCGAGATCCAA	CAATCTCGCTGGCGGTCAAGAAAGGACGACCGCACGGCC	1080		
1097	TACTGGGCGAGATCCAA	CAATCTCGCTGGCGGTCAAGAAAGGACGACCGCACGGCC	1156		
1081	ACCGGACCCCTGTGGT	TATTCGCGCATGACGCGCGCTTACGCTTGAGAGTGTGACCGCG	1140		
1157	ACCGGACCCCTGTGGT	TATTCGCGCATGACGCGCGCTTACGCTTGAGAGTGTGACCGCG	1216		
1141	GGCCCCCTGCGGGA	CGGGTGGGGAAGAGGGAGCGCGGCTGGGGA	1187		

```

Db      1217  GSCCCCTGCGGGACGGGTGGCGGAAGAGGAGGAGCGCGCTGGGGA 1263
|||||
AP001178          152711 bp   DNA       linear    PRI 11-JUL-2002
Homo sapiens genomic DNA, chromosome 18 clone:RP11-806L2, complete
sequence.
ACCESSION      AP001178          GI:21728142
KEYWORDS       HTG.
SOURCE         Homo sapiens DNA, clone:RP11-806L2.
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1
AUTHORS        Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE          Homo sapiens genomic DNA
JOURNAL        Published Only in Database (2000)
REFERENCE      2 (bases 1 to 152711)
AUTHORS        Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
DIRECT SUBMISSION
TITLE          Direct Submission
JOURNAL        Submitted (14-FEB-2000) Masahira Hattori. The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT        On Jul 10, 2002 this sequence version replaced gi:9229948.
FEATURES             Location/Qualifiers
     source           1..152711
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="18"
                     /map="18p"
                     /clone="RP11-806L2"
BASE COUNT      42612 a 32833 c 33397 g 43869 t
ORIGIN
Query Match      96.9%; Score 1149.8; DB 9; Length 152711;
Best Local Similarity 98.9%; Pred. No. 4.3e-138;
Matches 1174; Conservative 0; Mismatches 10; Indels 3; Gaps 2
QY      1      GATCGGCCACTGCATCCAGCCTGGGTGAGAGCGAGCATCTGTCTCAAAAAAAAAA 60
Db      84508  GATCGCACCAATTGCATCCAGCCTGGGTGAGAGCGAGCATCTGTCTCAAAAAAAAAA 84567
QY      61      AAAAAGACCGCCAGGGCTCAAACAAAAAACTCGGAAAAGCCCTGGCGTCTTTTTTTT 120
Db      84568  AAAAAGACCGCCAGGGCTCAAACAAAAAACTCGGAAAAGCCCTGGCGTCTTTTTTTT 84625
QY      121     TTTTCTTTTTTTTTTTTTTTGGGACAGTCCTTGCTCTGTGCCACAGCTGGAGTAGTAATGT 180
Db      84626  TTTTTTTTTTTTTTTTTTTTTTTGGGACAGTCCTTGCTCTGTGCCACAGCTGGAGTAGTAATGT 84685
QY      181     CGGATCTTGCTCACTGCGAACCTCTGCTCTGCCCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 240
Db      84686  CGGATCTTGCTCACTGCGAACCTCTGCTCTGCCCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 84745
QY      241     TCCCAGTAGCCACACGCCAGCTAAATTTTGTANTTTTAGTAGACACGGGGTTTCAC 300
Db      84746  TCCCAGTAGCCACACGCCAGCTAAATTTTGTANTTTTAGTAGACACGGGGTTTCAC 84805
QY      301     CATGTTGTCCAGSGTGGTCTNGAACTCTGACCTCAGGTGATCCACCGGCTCGGCCCCC 360
Db      84806  CATGTTGTCCAGSGTGGTCTTGAATCTCTGACCTCAGGTGATCCACCGGCTCGGCCCCC 84865
QY      361     CAAAGTACTAGGATTACAGCGGTGAGCCACCGCGTCCAGCGCCCTGGCGGTTTAAATCA 420
Db      84866  CAAAGTACTAGGATTACAGCGGTGAGCCACCGCGTCCAGCGCCCTGGCGGTTTAAATCA 84925

```

Q17 1141 GGCCCTTGGCGGGA CCGGTGGCGGGA AGGAGGCGCGGCTGGGGA 1187

Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
Dearellano, K., Dewar, K., Domino, M., Doyle, M., Penestor, J.,
Perreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,
Gardina, S., Grant, G., Hagos, B., Hearford, A., Horton, L.,
Howland, J., Johnson, R., Jones, C., Kann, L., Karakas, A., Klein, J.,
Landers, T., Lehotsky, J., Levine, R., Liu, C., Liu, G., Locke, K.,
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
McPheters, R., Meldrum, J., Menes, L., Morrow, J., Naylor, J.,
Norman, C.H., O'Connor, T., O'Donnell, P., Oliver, T.M., Peterson, K.,
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,
Zimmer, A. and Zody, M.

Direct Submission
Submitted (08-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 1, 2000 this sequence version replaced gi:6751720.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information
Center project name: L976
Center clone name: 145_B_19
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 160895 bases at least Q40
Consensus quality: 163157 bases at least Q30
Consensus quality: 164107 bases at least Q20
Insert size: 164000; agarose-fp
Insert size: 164942; sum-of-contigs
Quality coverage: 5.9 in Q20 bases; agarose-fp
Quality coverage: 5.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1339: contig of 1339 bp in length
1340 1439: gap of 100 bp
1440 9151: contig of 7712 bp in length
9152 9251: gap of 100 bp
9252 13650: contig of 10399 bp in length
13651 13750: gap of 100 bp
13751 34038: contig of 14288 bp in length
34039 34138: gap of 100 bp
34139 47814: contig of 13676 bp in length
47815 47914: gap of 100 bp
47915 62653: contig of 14739 bp in length
62654 62753: gap of 100 bp
62754 89031: contig of 26278 bp in length
89032 89131: gap of 100 bp
89132 116608: contig of 27475 bp in length
116609 116706: gap of 100 bp
116707 165742: contig of 49036 bp in length.

Location/Qualifiers
1..165742
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="RP11"
/clone="RP11-145B19"

TITLE
JOURNAL
COMMENT
FEATURES
source

QY 182 GGATTTGGTCTGCAACCTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCCT 241
Db 1005 GAACTTTAACTCTACTACCACTCTACTCCCAAAATTCACCAATCTTCTACCTCAACCT 946
QY 242 CCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTTAGTAGAGAGGGGTTTCAAC 301
Db 945 CCCAAATAACCAACCGCCCACTAAATTTTATATCTTTTAAATAAAGAAATTTCAAC 886
QY 302 ATGTTGTCAGGCTGTCTGACCTGACCTCAGGTGATCCACCGCTCGGCCCCC 361
Db 885 ATATTATCCAACTAATCTTAACTCTTAACCTCAAAATATCCACCGCTCGGCCCCC 826
QY 362 AAAGTACTAGGATTACAGCGGTGAGCCACCGCTCCAGCGCTCGCGGTTTAAATCAA 421
Db 825 AAAATATCTAAATTAACAGGTAAACACCGCTCCAGCGCTCAAGATTTTAAATCAA 766
QY 422 GTAGAAAGCTGCAATATACCACTTGTCTTCGTTTCGTTTCAGTGAGAGAAATGGA 481
Db 765 ATAAATAAATCAATATATACCACTTACTTCGATTAC-TTCAATAAATAAGAAATAAA 707
QY 482 AATGCAATCMTCTTATTTAGTTGTAGAAACAGATCTCAACAGAGTTTGTGNGACAGA 541
Db 706 AATACAAATCCCTTATTAATATAAATAAACAATCTCAACAACAATTTTATTAACAAA 647
QY 542 CGCAGGAAACCGTGGAACTGTGTCTGCTGTGAGTGTAGAGAGGCGCGTGCACCAACGCT 601
Db 646 CGCAGGAAACCGTGGAACTTACTACTTAAATAAATAAAGCGGATCGACCAACGAT 587
QY 602 TCCAAAGGCGAGTCTCTCCGNGCCACCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
Db 586 TCCAAAGGCGAGTCTCTCCGNGCCACCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527
QY 662 AGACTCTCAGTGTGGCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 721
Db 526 AAACCTCAATATAACCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 467
QY 722 CCCTGGCGCAGCTCTAGAGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 781
Db 466 CCCTAACGCACTCTCTTAACAGAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 407
QY 782 AGCGCGGACCGCGCGGAGGCGCGCGGAGGCGCGCGGAGGCGCGCGGAGGCGCGGAGGCG 841
Db 406 AACGGAACGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 347
QY 842 CTGCTCTCGCTCCCGCGCGCGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 901
Db 346 CTACCTCGCTCCCGCGCGCGCACTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 287
QY 902 CTGCTGCTCCCGCGCGCGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 961
Db 286 CTGCTGCTCCCGCGCGCGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 227
QY 962 TGCCCG 1021
Db 226 TACCCCG 167
QY 1022 ACCCTGGGCGAGTCAACATCTCTGCTGCGGCGGTGAGGAGGAGGAGGAGGAGGAGGAGG 1081
Db 166 ACCTAAACAAATCAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 107
QY 1082 CCGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1141
Db 106 CCGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6
QY 1142 GCGCTGCGGCGCGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1182
Db 46 ACCCTACGAAACGAAATACGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6

RESULT 13
AX348476 1186 bp DNA linear PAT 06-FEB-2002
LOCUS
DEFINITION Sequence 171 from Patent WO0202806.

AX348476
AX348476.1 GI:18614512
synthetic construct.
synthetic construct
artificial sequences.
REFERENCE
Olek, A., Piepenbrock, C. and Berlin, K.
Method and nucleic acids for pharmacogenomic methylation analysis
Patent: WO 0202806-A 171 10-JAN-2002;
Epigenomics AG (DE)
FEATURES
Location/Qualifiers
1..1186
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 224 a 98 c 347 g 517 t
Query Match 59.8%; Score 710; DB 6; Length 1186;
Best Local Similarity 75.1%; Pred. No. 1.4e-118;
Matches 892; Conservative 0; Mismatches 294; Indels 1; Gaps 1;
QY 1 GATCGCGCACTGCATCCAGCCTGGTGCAGAGCGAGACTCTGTCTCAAAAAAAA 60
Db 1 GATCGCGTTATTTGATTTTGTGTTGGTGCAGAGCGAGATTTGTTTAAAAAAA 60
QY 61 AAAAAAGCGCGAGGCTCAAAACAAAAACCTCGGAAAGCCCTGGCGGCTCTTTT 120
Db 61 AAAAAAGCTTAGGTTTAAATAAAAAATTCGGAAGTTTGGCGGTTT 120
QY 121 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 180
Db 121 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 180
QY 181 CGSACTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 240
Db 181 CGSACTTGGCTTATGTAATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 240
QY 241 TCCCAAGTAGCCACCAAGCGGCTAATTTTGTANTTTTAGTAGAGAGGGGTTTCA 300
Db 241 TTTTAAAGTAGTTATTTAGCTTTAGTTAATTTTGTATTTTAGTAGAGAGGGGTTTAT 300
QY 301 CATGTTGCCAGCTGTCTGAGACTCTGATCCAGCTGATCCACCGCTCGGCCCCC 360
Db 301 TATGTTGTTAGTTGTTGTTTGAATTTTCAATTTTAGGTGATTTATTCGTTTCGTT 360
QY 361 CAAAGTACAGGATTAAGCGGTGAGCCACCGGTTCCAGCGCTCGCGGCTTTTAAATCA 420
Db 361 TAAAGTATTTAGGATTTATAGCGTGAGTTATCGCGTTTGGCGGTTTGGCGGTTTAAATTA 420
QY 421 AGTAGAAGAGCTGCAATATATACCACTTCTGCTTCGTTGCTTCAAGTGAGAGCAAGAAATGG 480
Db 421 AGTAGAAGAGTTGTAATTAATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 479
QY 481 AAATGCAATCMTCTTATTTAGTTGTAGGAAACAGATCTCAACAGCAGTCTTGTGNGACAAG 540
Db 480 AAATGTAATTTTATTTAGTTGTAGGAAATAGATTTTAAATAGTAGTTTGTGTTGATAAG 539
QY 541 ACCGAGGAAAAAGTGGGAACCTGTGCTGTGCTGTAGAGAGGCGCGGTGCACAGACGG 600
Db 540 ATCGTAGGAAACGTTGGGAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 599
QY 601 TTCCCAAGGCGAGCTCCTTCCCGCCACCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 600 TTTTAAAGGCGTAGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 659
QY 661 AAGACTCTCAGCTGTGCGCTTGGCTCCGTTCTGTGCGCACACCCGCTGCTGCTGCTGCTGCT 720
Db 660 AAGATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTT 719
QY 721 CCCCTGGCGCAGCTCTCTAGAGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780

b 720 TTTTGGGCTACGTTTATTTAGACGGGGGTGTCGCGATTTCGTCGAGTAGGAAGAGCG 779
y 781 GAGCGGGGACGGCCGCGGAAAGCGCGGGAAGGGTCTTCCACCGCGCCACTTGG 840
b 780 GAGCGGGGACGGTCGCGGAAAGCGCGGGAAGGGTCTTCCACCGCGCCACTTGG 839
y 841 CTGCTCCGTCG 900
b 840 TTTGTTTTCGTTTCGTCGCGCTATTTTGGTTTTCGTTTCGTCGCGCTATTTTCGTTG 899
y 901 CCTCGTCCCG 960
b 900 TTTTCGTTTTCGTCGCGCTATTTTGGTTTTCGTTTCGTCGCGCTATTTTCGTTG 959
y 961 TTGCGCCCG 1020
b 960 TTGTTTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1019
y 1021 TACTGGGCGAGATCAACACATCTTCCTCGCTCGCGGTCGAGAAAGGACGACCGCG 1080
b 1020 TATTTGGGTAGATTAATAATATTTTTCGTTTCGCGCTTAGGAAGGACGATCGTGGGT 1079
y 1081 ACCGCGACCTGTCGTTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1140
b 1080 ATCGGTATTTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1139
y 1141 GCGCCCTCGCGGACGGTGGCGGAAGGAGGAGCGCGCGCTGGGGA 1187
b 1140 GGTTCGCGGACGGGTGGCGGAAGGAGGAGCGCGCGCTGGGGA 1186

RESULT 14
LOCUS AR103889 1536 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 1 from patent US 6087489.
ACCESSION AR103889
VERSION AR103889.1 GI:12815477
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1536)
AUTHORS Dean,N.M.
TITLE Antisense oligonucleotide modulation of human thymidylate synthase expression
JOURNAL Patent: US 6087489-A 1 11-JUL-2000;
FEATURES Location/Qualifiers
source i. .1536
/organism="unknown"
BASE COUNT 390 a 369 c 399 g 378 t
ORIGIN

Query Match 25.2%; Score 299.4; DB 6; Length 1536;
Best Local Similarity 99.7%; Pred. No. 1.8e-44;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 833 CCACCTTGGCTGCTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 892
DB 14 CCACCTTGGCTGCTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 73
QY 893 TTCGCTGCTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 952
DB 74 TTCGCTGCTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 133
QY 953 GCGGCGCTTGGCG 1012
DB 134 GCGGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 193
QY 1013 AGCTCAGTACTGGGCGAGATCCAAACATCTTCGCTCGCGGCTCAGGAAGGACG 1072
DB 194 AGCTCAGTACTGGGCGAGATCCAAACATCTTCGCTCGCGGCTCAGGAAGGACG 253
QY 1073 GCACGGGCGACCGGACCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 1132
DB 254 GCACGGGCGACCGGACCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 313

QY 1133 A 1133
DB 314 A 314

Search completed: November 28, 2002, 21:45:16
Job time : 3331.72 secs

Db 254 GCACGGGCGACCGGACCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 313
QY 1133 A 1133
DB 314 A 314

RESULT 15
LOCUS AX397715 1536 bp DNA linear PAT 18-MAY-2002
DEFINITION Sequence 141 from Patent WO0210366.
ACCESSION AX397715
VERSION AX397715.1 GI:21068437
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Chen,H., Preimer,N.B. and Novak,T.
TITLE Methods and compositions for diagnosing and treating chromosome-18P related disorders
JOURNAL Patent: WO 0210366-A 141 07-FEB-2002;
Millennium Pharmaceuticals, Inc. (US) ; The Regents of The University of California (US)
FEATURES Location/Qualifiers
source i. .1536
/organism="Homo sapiens"
/db xref="taxon:9606"
BASE COUNT 390 a 369 c 399 g 378 t
ORIGIN

Query Match 25.2%; Score 299.4; DB 6; Length 1536;
Best Local Similarity 99.7%; Pred. No. 1.8e-44;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 833 CCACCTTGGCTGCTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 892
DB 14 CCACCTTGGCTGCTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 73
QY 893 TTCGCTGCTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 952
DB 74 TTCGCTGCTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 133
QY 953 GCGGCGCTTGGCG 1012
DB 134 GCGGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 193
QY 1013 AGCTCAGTACTGGGCGAGATCCAAACATCTTCGCTCGCGGCTCAGGAAGGACG 1072
DB 194 AGCTCAGTACTGGGCGAGATCCAAACATCTTCGCTCGCGGCTCAGGAAGGACG 253
QY 1073 GCACGGGCGACCGGACCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 1132
DB 254 GCACGGGCGACCGGACCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 313

QY 1133 A 1133
DB 314 A 314

Search completed: November 28, 2002, 21:45:16
Job time : 3331.72 secs

BASE COUNT 217 a 396 c 259 g 188 t

ORIGIN

Query Match 28.2%; Score 335; DB 13; Length 1060;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 795 CGCGGAAAGCGCGGAGGGGTCTGCGACCGCGCCTTGGCTGCTCCGTCCTCC 854
 Db 1 CGCGGAAAGCGCGGAGGGGTCTGCGACCGCGCCTTGGCTGCTCCGTCCTCC 60

Qy 855 GCGCGCCACTTGGCTGCTCCGTCCTGCGCGCGCGCCTTGGCTGCTCCGTCCTCC 914
 Db 61 GCGCGCCACTTGGCTGCTCCGTCCTGCGCGCGCGCCTTGGCTGCTCCGTCCTCC 120

Qy 915 CGCGCGCGCATGCTGCTGCGCGCGCTCGAGTGTGCGCGCGCGCTTGGCTGCTCCGTCCTCC 974
 Db 121 CGCGCGCGCATGCTGCTGCGCGCGCTCGAGTGTGCGCGCGCGCTTGGCTGCTCCGTCCTCC 180

Qy 975 ACAGGACGGGACGCGCGCGCTGCGCGCGCTCGAGTGTGCGCGCGCGCTTGGCTGCTCCGTCCTCC 1034
 Db 181 ACAGGACGGGACGCGCGCGCTGCGCGCGCTCGAGTGTGCGCGCGCGCTTGGCTGCTCCGTCCTCC 240

Qy 1035 CCAACACATCTCCGCTGCGCGCTCAGAGAGACGCGCGCGCGCTTGGCTGCTCCGTCCTCC 1094
 Db 241 CCAACACATCTCCGCTGCGCGCTCAGAGAGACGCGCGCGCGCTTGGCTGCTCCGTCCTCC 300

Qy 1095 GGTATTGGGATGCGGCGCGCTACAGCTGAGAG 1129
 Db 301 GGTATTGGGATGCGGCGCGCTACAGCTGAGAG 335

RESULT 2

AU118644

LOCUS AU118644 HEMBA1 Homo sapiens cDNA clone HEMBA1004057 5', mRNA

DEFINITION sequence.

ACCESSION AU118644.1 GI:10933771

VERSION AU118644.1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE 1 (bases 1 to 741)

COMMENT Ota.T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.

TITLE HRI human cDNA project

JOURNAL Unpublished (2000)

COMMENT Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomics@hri.co.jp

HRI human cDNA project, 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

Location/Qualifiers

source 1..741

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="HEMBA1004057"

/clone_lib="NIH_MGC_85"

/tissue_type="lymphoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: lymph; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."

BASE COUNT 153 a 227 c 218 g 140 t

ORIGIN

Query Match 27.7%; Score 329; DB 9; Length 741;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 801 AAAAGCGCGCGGAGGGGTCTGCGACCGCGCCTTGGCTGCTCCGTCCTCCGTCCTCC 860
 Db 2 AAAAGCGCGCGGAGGGGTCTGCGACCGCGCCTTGGCTGCTCCGTCCTCCGTCCTCC 61

Qy 861 CCATTGGCTGCTGCTGCTGCGCGCGCTCGAGTGTGCGCGCGCGCTTGGCTGCTCCGTCCTCC 920
 Db 62 CCATTGGCTGCTGCTGCTGCGCGCGCTCGAGTGTGCGCGCGCGCTTGGCTGCTCCGTCCTCC 121

Qy 921 CGCCATGCTGCTGCTGCTGCGCGCGCTCGAGTGTGCGCGCGCGCTTGGCTGCTCCGTCCTCC 980
 Db 122 CGCCATGCTGCTGCTGCTGCGCGCGCTCGAGTGTGCGCGCGCGCTTGGCTGCTCCGTCCTCC 181

Qy 981 GCGGAGCGCGCGCGCTGCTGCGCGCGCTCGAGTGTGCGCGCGCGCTTGGCTGCTCCGTCCTCC 1040
 Db 182 GCGGAGCGCGCGCGCTGCTGCGCGCGCTCGAGTGTGCGCGCGCGCTTGGCTGCTCCGTCCTCC 241

Qy 1041 CATCTCCGCTGCGCGCTCAGAGAGACGCGCGCGCGCTTGGCTGCTCCGTCCTCCGTCCTCC 1100
 Db 242 CATCTCCGCTGCGCGCTCAGAGAGACGCGCGCGCGCTTGGCTGCTCCGTCCTCCGTCCTCC 301

Qy 1101 CGGCATGCGGCGCGCTACAGCTGAGAG 1129
 Db 302 CGGCATGCGGCGCGCTACAGCTGAGAG 330

RESULT 3

BQ434233

LOCUS BQ434233 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6159722

DEFINITION 5', mRNA sequence.

ACCESSION BQ434233

VERSION BQ434233.1 GI:21173309

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE 1 (bases 1 to 841)

COMMENT NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTP/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

TITLE DNA Sequencing by: Agencourt Bioscience Corporation

JOURNAL DNA distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLAM13509 row: e column: 03
 High quality sequence stop: 639.

Location/Qualifiers

source 1..841

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6159722"

/clone_lib="NIH_MGC_72"

/lab host="DH10B (phage-resistant)"
 /note="Organ: melanoma"
 /lab host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pcwv-sport6; Site: 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

BASE COUNT 176 a 263 c 241 g 161 t
 ORIGIN
 Query Match 26.8%; Score 318; DB 14; Length 841;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 812 GGAAGGCTCTGCAACCGGCGCACTTGGCTGCTCCGTCGCGCGCGCGCACTTGGCT 871
 Db 1 GGAAGGCTCTGCAACCGGCGCACTTGGCTGCTCCGTCGCGCGCGCGCACTTGGCT 60
 QY 872 GCTTCCGTCGCGCGCGCGCACTTGGCTGCTCCGTCGCGCGCGCGCGCACTTGGCT 931
 Db 61 GCTTCCGTCGCGCGCGCGCACTTGGCTGCTCCGTCGCGCGCGCGCGCACTTGGCT 120
 QY 932 TGGCGGCTGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 991
 Db 121 TGGCGGCTGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 QY 992 AGCGGCTGCG 1051
 Db 181 AGCGGCTGCG 240
 QY 1052 GCGGCTGCGAGGAGCG 1111
 Db 241 GCGGCTGCGAGGAGCG 300
 QY 1112 CGCGTACAGCTGAGAG 1129
 Db 301 CGCGTACAGCTGAGAG 318

RESULT 4
 BQ685120 1021 bp mRNA linear EST 15-JUL-2002
 LOCUS AGENCOURT 8344472 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6251301
 5', mRNA sequence.

ACCESSION BQ685120

VERSION BQ685120.1 GI:21810436

KEYWORDS EST.

SOURCE human.

ORGANISM

REFERENCE
 1 (bases 1 to 1021)
 NIH-MGC http://mgi.nci.nih.gov/

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: c9apb8-remail.nih.gov

Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov

plate: LLCW2395 row: 1 column: 22

High quality sequence start: 12

High quality sequence stop: 409.

Location/Qualifiers

1..1021

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6251301"

/clone_lib="NIH_MGC_110"

/tissue_type="ductal carcinoma, cell line"

FEATURES

source

/lab host="DH10B (phage-resistant)"
 /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."

BASE COUNT 202 a 335 c 289 g 194 t

ORIGIN

Query Match 26.0%; Score 309; DB 14; Length 1021;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 309; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 821 CTGTCACCGCGCACTTGGCTGCTCCGTCGCGCGCGCGCACTTGGCTGCTCCGTC 880
 Db 11 CTGTCACCGCGCACTTGGCTGCTCCGTCGCGCGCGCGCGCACTTGGCTGCTCCGTC 70
 QY 881 CGCGCGCGCACTTGGCTGCTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 940
 Db 71 CGCGCGCGCACTTGGCTGCTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 130
 QY 941 CGGAGCTGCG 1000
 Db 131 CGGAGCTGCG 190
 QY 1001 CG 1060
 Db 191 CG 250
 QY 1061 GGAAGGACG 1120
 Db 251 GGAAGGACG 310
 QY 1121 GCTTGAGAG 1129
 Db 311 GCTTGAGAG 319

RESULT 5

AL551990

LOCUS AL551990 LTI_NFL006 PL2 Homo sapiens cDNA clone CS01060YE04 5

DEFINITION prime, mRNA sequence.

ACCESSION AL551990

VERSION AL551990.1 GI:12890459

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 938)

Li W.B., Gruber, C., Jesse, J., and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1..938

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS01060YE04"

/clone_lib="LTI_NFL006_PL2"

/tissue_type="placenta"

/note="Vector: pcwvSPORT 6; Site 1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pcwvSPORT 6
 vector. Library was normalized. Library was constructed by

Life Technologies. Contact : Peng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 202 a 288 c 259 g 188 t 1 others
ORIGIN

Query Match 25.5%; Score 303; DB 9; Length 938;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 827 ACCGCGCCACTTGGCTCCCTCCCGCGCGCCACTTGGCTCCGCTCCGTCGCCGCG 886

Db 1 ACCGCGCCACTTGGCTCCCTCCCGCGCGCCACTTGGCTCCGCTCCGTCGCCGCG 60

QY 887 CGCACCTTGGCTCCCTCCCGCGCGCCACTTGGCTCCGCTCCGTCGCCGCG 946

Db 61 CGCACCTTGGCTCCCTCCCGCGCGCCACTTGGCTCCGCTCCGTCGCCGCG 120

QY 947 TGGCGCGCGCGCTTGGCTCCCGCGCGCCACTTGGCTCCGCTCCGTCGCCGCG 1006

Db 121 TGGCGCGCGCGCTTGGCTCCCGCGCGCCACTTGGCTCCGCTCCGTCGCCGCG 180

QY 1007 ACGGGAGCTGAGTACCTGGGGGAGATCCACACATCTCCGTCGGGGCTCAGGAAG 1066

Db 181 ACGGGAGCTGAGTACCTGGGGGAGATCCACACATCTCCGTCGGGGCTCAGGAAG 240

QY 1067 ACGACGCGACGGGACCGGACCTGTCGGTATTCGGCATGACGGCGCTACAGCTGA 1126

Db 241 ACGACGCGACGGGACCGGACCTGTCGGTATTCGGCATGACGGCGCTACAGCTGA 300

QY 1127 GAG 1129

Db 301 GAG 303

RESULT 6

AL541351 AL541351 LTI FL002_PL1 Homo sapiens cDNA clone CS0E006YA08 5 prime
DEFINITION
mRNA sequence.

ACCESSION AL541351.1 GI:12872339

VERSION EST.

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 668)

AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. 668

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0E006YA08"

/clone_lib="LTI FL002_PL1"

/lab_host="DH10B"

/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand

cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-stranded cDNA was digested with Not I

and cloned into the Not I and Eco RV sites of the

pCMVSPORT 6 vector. Library was constructed by Life

Technologies. Contact : Peng Liang Life Technologies, a

division of Invitrogen 9800 Medical Center Drive Rockville

, Maryland 20850, USA Fax : (1) 301 610 8371 Email :

fliang@lifetech.com URL :

BASE COUNT 139 a 208 c 192 g 126 t 3 others
ORIGIN

Query Match 25.2%; Score 299; DB 9; Length 668;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 820 TCTTGGCCACCGCCCACTTGGCTCCGCTCCGTCGCCGCGCCACTTGGCTCCGCTCCGTC 879

Db 1 TCTTGGCCACCGCCCACTTGGCTCCGCTCCGTCGCCGCGCCACTTGGCTCCGCTCCGTC 60

QY 880 CCCCGCGCGCCACTTGGCTCCGCTCCGTCGCCGCGCGCCACTTGGCTCCGCTCCGTC 939

Db 61 CCCCGCGCGCCACTTGGCTCCGCTCCGTCGCCGCGCGCCACTTGGCTCCGCTCCGTC 120

QY 940 TCGAGCTCCGCGCGCGCCCTTGGCTCCGTCGCCGCGCGCCACTTGGCTCCGCTCCGTC 999

Db 121 TCGAGCTCCGCGCGCGCCCTTGGCTCCGTCGCCGCGCGCCACTTGGCTCCGCTCCGTC 180

QY 1000 CCGCGCGACGGGAGCTGAGTACCTGGGCGACATCCAAACATCTCCGTCGCCGCGCTA 1059

Db 181 CCGCGCGACGGGAGCTGAGTACCTGGGCGACATCCAAACATCTCCGTCGCCGCGCTA 240

QY 1060 AGGAAGGACGACGCGACCGGCGACCGCTGTCGGTATTCGGCATGACGGCGCGCTA 1118

Db 241 AGGAAGGACGACGCGACCGGCGACCGCTGTCGGTATTCGGCATGACGGCGCGCTA 299

RESULT 7

BM763658

LOCUS

DEFINITION

mRNA sequence.

ACCESSION

BM763658

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 523)

AUTHORS Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,

Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and

Kim, Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished (2002)

COMMENT Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 14 row: F column: 09

High quality sequence stop: 523.

Location/Qualifiers

1. 523

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="S13KMS5-14-F09"

/clone_lib="S13KMS5"

/tissue_type="myeloma"

/cell_line="KMS-5"

/lab_host="Top10F"

/note="Vector: pCNS; Site 1: EcoRI; Site 2: NotI; The poly

(A) + RNA was dephosphorylated with bacterial alkaline

phosphatase (BAP) and then decapped with tobacco acid

pyrophosphatase (TAP). The decapped intact mRNA was

ligated with DNA-RNA linker including EcoR I site by

treatment of T4 RNA ligase and the first strand cDNA was

synthesized from oligo dT-selected mRNA by priming with

dT-tailed vector. The dT-tailed vector was adjusted to

this is a NIH MGC Library."

BASE COUNT	208 a	287 c	261 g	192 t
ORIGIN				

ORIGIN

Query Match 23.6%; Score 280; DB 14; Length 948;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 280; Conservative 0; Mismatches 0; Indels

Qy	850	GTCCGCGCGGCGCAATTGGCTGCTCCGTCGCGCGCACTTCGCTGCTCCGTC	909
Db	1	GTCCGCGCGCGCACTTCGCTGCTCCGTCGCGCGCGCACTTCGCTGCTCCGTC	60
Qy	910	CCCGCCGCGCGCCATCTGTGGCGCGCTCGAGTTCGCGCGCGCCCTTGTCCCCCC	969
Db	61	CCCGCCGCGCGCGCCATCTGTGGCGCGCTCGAGTTCGCGCGCGCCCTTGTCCCCCC	120
Qy	970	GCGCACAGAGCGGAGCGCGAGCGCGTCCGCGCGACGCGGAGCTGCATCTCTGGGG	1029
Db	121	GCGCACAGAGCGGAGCGCGAGCGCGTCCGCGCGACGCGGAGCTGCATCTCTGGGG	180
Qy	1030	CGAGTCCAAACACATCTCCGTGCGGCGTTCAGGAGGACACCGCACCGGCGCACCGGCACC	1089
Db	181	CGAGTCCAAACATCTCCGTGCGGCGTTCAGGAGGACACCGCACCGGCGCACCGGCACC	240
Qy	1090	CTGTTCGATTTGGCGATGACGGCGCGTCAAGCTTGAGAG	1129
Db	241	CTGTTCGATTTGGCGATGACGGCGCGTCAAGCTTGAGAG	280

RESULT 10	linear	EST 05-AUG-2002
AU143295	739 bp	mRNA
LOCUS	AU143295	V78A1 Homo sapiens cDNA clone Y79A1001699 5', mRNA

DEFINITION	
SOURCE	human
ORGANISM	Homo sapiens
KEYWORDS	Cranialata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata;
VERSION	AUI43295.1 GI:11004816
ACCESSION	AUI43295
SEQUENCE	
DESCRIPTION	
REMARKS	
FEATURES	
REFERENCES	
COMMENTS	
DATE	1997-07-01
ENTRY	
END	

REFERENCE
1 (bases 1 to 739)
Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
Yanamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and
Tsurai, T.

TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomics@hri.co.jp
 HRI human cDNA project; 5' & 3'-end one pass sequencing; Helix
 Research Institute; cDNA library construction; Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute

```

FEATURES
source
helix_residues
location/Qualifiers
1..739
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Y79AA1001699"
/clone_lib="vY9AA1"
/cell_type="retinoblastoma"
/cell_line="Y79"
/note="Vector: pME18SFL3"

```

Query Match	23.4%;	Score 278;	DB 9;	Length 739;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 328.	Conservative	0:	Mismatches 1:	Indels 0:
Gaps	0:			

Qy	801	AAAAGGCGCGGAGAGGGGTCTGCGCACCGCGGCACATTGGCTTGCTCCGCTCCGCGCGG 860
Db	2	AAAAGGCGCGGAGAGGGGTCTGCGCACCGCGGCACATTGGCTTGCTCCGCTCCGCGCGG 61
Qy	861	CCACTTGGCGCTGCTCCGTCGCCGCGCGGCACATTGGCTTGCTCCGCTCCGCGCGCGCGG 920
Db	62	CCACTTGGCGCTGCTCCGTCGCCGCGCGGCACATTGGCTTGCTCCGCTCCGCGCGCGCGG 121
Qy	921	CGCCATGCTGTGCGCGGCTCGGAGCTGCGCGCGCGCGCCTTGGCCCGCGCGCAGAGA 980
Db	122	CGCCATGCTGTGCGCGGCTCGGAGCTGCGCGCGCGCCTTGGCCCGCGCGCAGAGA 181
Qy	981	GCGGAGCGCGGAGCCGCGCTCCGCCGACGCGGAGCTGCAGTACCTTGGGCGAGATCCAA 1040
Db	182	GCGGAGCGCGGAGCCGCGGTCCGCCGACGCGGAGCTGCAGTACCTTGGGCGAGATCCAA 241
Qy	1041	CATCTTCGCTGCGGGCTCAGGAAGACGACCGGACCGGACACCTGTGCGTATT 1100
Db	242	CATCTTCGCTGCGGGCTCAGGAAGACGACCGGACCGGACACCTGTGCGTATT 301
Qy	1101	CGGATGCGGGCGCGTACAGCTGAGAG 1129
Db	302	CGGATGCGGGCGCGTACAGCTGAGAG 330

RESULT 11	819 bp	mRNA	linear	EST 05-AUG-2002
AU143180	AU143180	v792a1	Homos sapiens cDNA clone v792a1001513	5' mRNA
LOCUS				

DEFINITION	AU143180.1 HUMAN HOMO SAPIENS COM1 GENE, INHIBITORS 3, / HUMAN
SEQUENCE	sequence.
ACCESSION	
VERSION	AU143180
KEYWORDS	AU143180.1 GI:11004701
EST.	EST.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Actinopterygii; Clupeiformes; Clupeidae; Harengidae; Homo

REFERENCE
1 (bases 1 to 819)
Cta, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and
Isogai, T.

HRI human cDNA project
 Unpublished (2000)
 Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3875
 Fax: 81-438-52-3986
 Email: genomics@hri.co.jp
 HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
 Research Institute; cDNA library construction: Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute

```

FEATURES
source
Location/Qualifiers
1. .819
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Y79AA1001513"
/clone_lib="Y79AA1"
/cell_type="retinoblastoma"
/cell_line="Y79"
/note="Vector: pME18SF13"
169 a 243 c 239 g 164 t
BASE COUNT
ORIGIN
4 others

```

Query Match 23.4%; Score 278; DB 9; Length 819;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 328; Conservative 0; Mismatches 1; Indels 0; Gaps 0

Qy 801 AAAAGCGCGGGAAGGGGTCTGTCCACCGCGCCACTTGGCTGCCTCGTCCGCGCGG 860

```

Db 2 AAAAGCGCGCGAAGGGTCTGTGCAACGCGCCACTTGGCTTCGCTCCGTCGCCGCG 61
Qy 861 CCACTTGGCTGCTCCTCCGCGCGCGCCACTTGGCTTCGCTCCGCTCCGCGCGCGCG 920
Db 62 CCACTTGGCTGCTCCTCCGCGCGCGCCACTTGGCTTCGCTCCGCTCCGCGCGCGCG 121
Qy 921 CGCATGCTGCTGGCGGCTCCGAGTCCGCGCGCGCGCTTGGCTCCGCGCGCGCGCA 980
Db 122 CGCATGCTGCTGGCGGCTCCGAGTCCGCGCGCGCGCTTGGCTCCGCGCGCGCGCA 181
Qy 981 CGGAGCGCGAGCGCGCTCCGCGCGCGCGAGTCCGCGCGCGCGCTTGGCTCCGCGCG 1040
Db 182 GGGGAGCGCGAGCGCGCTCCGCGCGCGCGAGTCCGCGCGCGCGCTTGGCTCCGCGCG 241
Qy 1041 CATCTCTCGCTCGCGGCTCAGGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1100
Db 242 CATCTCTCGCTCGCGGCTCAGGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 301
Qy 1101 CGGCATGCGAGCGCGCTTACAGCTGAGAG 1129
Db 302 CGGCATGCGAGCGCGCTTACAGCTGAGAG 330

RESULT 12
BM478045 1045 bp mRNA linear EST 05-FEB-2002
LOCUS BM478045
DEFINITION AGENCOURT 6482920 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5552680
5', mRNA sequence.
ACCESSION BM478045
VERSION BM478045.1 GI:18527087
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1045)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12268 row: 9 column: 17
High quality sequence stop: 687.
Location/Qualifiers
1..1045
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5552680"
/clone_lib="NIH MGC 85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: PCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."
BASE COUNT 242 a 297 c 286 g 220 t
ORIGIN
Query Match 23.4%; Score 278; DB 13; Length 1045;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 852 CCGCGCGCGCACTTGGCTGCTCGCTCCGCGCGCGCACTTGGCTGCTCGCTCCGCTCC 911
Db 1 CCGCGCGCGCACTTGGCTGCTCGCTCCGCGCGCGCACTTGGCTGCTCGCTCCGCTCC 60

```

```

Qy 912 CGCGCGCGCGCATGCTTGTGCGCGGCTCGAGTCTCGCGCGCGCGCTTGTGCGCGCGCG 971
Db 61 CGCGCGCGCGCGCATGCTTGTGCGCGGCTCGAGTCTCGAGTCTCGCGCGCGCGCTTGTGCGCGCGCG 120
Qy 972 CGCACAGAGCGGCGAGCGCGCGCGCTCGCGCGCGCGAGCTGAGTACTTGGGGCA 1031
Db 121 CGCACAGAGCGGCGAGCGCGCGCGCTCGCGCGCGCGAGCTGAGTACTTGGGGCA 180
Qy 1032 GATCCAAACACATCTCTCCGCTCGCGCGCTCAGGAAGGACGACCGGACCGGCGACCCCT 1091
Db 181 GATCCAAACACATCTCTCCGCTCGCGCGCTCAGGAAGGACGACCGGACCGGCGACCCCT 240
Qy 1092 GTCGCTATTGCGCATGCGAGCGCGCTTACAGCTGAGAG 1129
Db 241 GTCGCTATTGCGCATGCGAGCGCGCTTACAGCTGAGAG 278

RESULT 13
BM424746 885 bp mRNA linear EST 23-MAY-2002
LOCUS BM424746
DEFINITION AGENCOURT 7912639 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6150060
5', mRNA sequence.
ACCESSION BM424746
VERSION BM424746.1 GI:21120061
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 885)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue procurement: ATCC/DCTP/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13484 row: b column: 13
High quality sequence stop: 528.
Location/Qualifiers
1..885
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6150060"
/clone_lib="NIH MGC 72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: PCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 177 a 287 c 257 g 164 t
ORIGIN
Query Match 23.3%; Score 276; DB 14; Length 885;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 854 CGCGCGCGCACTTGGCTGCTCGCTCCGCGCGCGCGCACTTGGCTGCTCGCTCCGCG 913
Db 48 CGCGCGCGCACTTGGCTGCTCGCTCCGCGCGCGCGCACTTGGCTGCTCGCTCCGCG 107
Qy 914 CGCGCGCGCGCATGCTTGTGCGCGCGCTCGGAGCTCGCGCGCGCGCTTGTGCGCGCGCG 973
Db 108 CGCGCGCGCGCATGCTTGTGCGCGCGCTCGGAGCTCGCGCGCGCGCTTGTGCGCGCGCG 167
Qy 974 CACAGAGAGCGGCGCGCGCGCGCTCGCGCGCGCGCTCGCGCGCGAGTGTGAGTACTTGGGGCAGA 1033

```


QY 1036 CAACACATCTCCGCTGCGGCGTCAGGAAGGACGACCGCACGGGACACCGGACACCTGTTCG 1095
Db |||||
QY 199 CAACACATCTCCGCTGCGGCGTCAGGAAGGACGACCGCACGGGACACCGGACACCTGTTCG 258
Db |||||
QY 1096 GTATTTCGGCATGCAAGCGCGGCTACAGCCTTGAGAG 1129
Db |||||
QY 259 GTATTTCGGCATGCAAGCGCGGCTACAGCCTTGAGAG 292
Db |||||

Search completed: November 29, 2002, 03:28:11
Job time : 1576.32 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

M nucleic - nucleic search, using sw model

run on: November 29, 2002, 00:40:45 ; Search time 676.542 Seconds
(without alignments)
1187.536 Million cell updates/sec

Title: US-09-963-333-7
erfect score: 1187
equence: 1 gatcgccactgcactcca.....gaggaggcgctg999ga 1187.

oring table: OLLIGO NUC
Gapop_60.0 , Gapext 60.0

searched: 344316 seqs, 338423730 residues

ord size : 15

total number of hits satisfying chosen parameters: 60992

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Listing first 45 summaries

- Database :
- 1: Pending Patents NA New.*
 - 2: /cgn2_6/ptodata/1/pna/PCT NEW COMB.seq.*
 - 1: /cgn2_6/ptodata/1/pna/US06 NEW COMB.seq.*
 - 3: /cgn2_6/ptodata/1/pna/US07 NEW COMB.seq.*
 - 4: /cgn2_6/ptodata/1/pna/US08 NEW COMB.seq.*
 - 5: /cgn2_6/ptodata/1/pna/US09 NEW COMB.seq.*
 - 6: /cgn2_6/ptodata/1/pna/US10 NEW COMB.seq.*
 - 7: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1187	100.0	1187	5	US-09-658-659B-7
2	811	68.3	18597	5	US-09-658-659B-8
3	301	25.4	1942	5	US-09-724-676-34523
4	301	25.4	1942	5	US-09-724-676A-34523
5	301	25.4	2042	5	US-09-724-676-34515
6	301	25.4	2042	5	US-09-724-676A-34515
7	301	25.4	2391	5	US-09-724-676-34597
8	301	25.4	2391	5	US-09-724-676A-34597
9	301	25.4	2411	5	US-09-724-676-34535
10	301	25.4	2411	5	US-09-724-676A-34535
11	301	25.4	2491	5	US-09-724-676-34589
12	301	25.4	2491	5	US-09-724-676A-34589
13	301	25.4	2527	5	US-09-724-676-34548
14	301	25.4	2527	5	US-09-724-676A-34548
15	301	25.4	3023	5	US-09-724-676-34565
16	301	25.4	3023	5	US-09-724-676A-34565
17	301	25.4	3123	5	US-09-724-676-34557
18	301	25.4	3123	5	US-09-724-676A-34557
19	301	25.4	3139	5	US-09-724-676-34581
20	301	25.4	3139	5	US-09-724-676A-34581
21	301	25.4	3239	5	US-09-724-676-34573
22	301	25.4	3239	5	US-09-724-676A-34573
23	297	25.0	1536	5	US-09-658-659B-6
24	262	22.1	3298	6	US-10-240-965-200
25	111	9.4	2083	5	US-09-724-676-34524
26	111	9.4	2083	5	US-09-724-676A-34524

c 27	111	9.4	2183	5	US-09-724-676-34516	Sequence 34516, A
c 28	111	9.4	2183	5	US-09-724-676A-34516	Sequence 34516, A
c 29	111	9.4	2552	5	US-09-724-676-34536	Sequence 34536, A
c 30	111	9.4	2552	5	US-09-724-676A-34536	Sequence 34536, A
c 31	111	9.4	2632	5	US-09-724-676-34590	Sequence 34590, A
c 32	111	9.4	2632	5	US-09-724-676A-34590	Sequence 34590, A
c 33	111	9.4	2652	5	US-09-724-676-34529	Sequence 34529, A
c 34	111	9.4	2652	5	US-09-724-676A-34529	Sequence 34529, A
c 35	111	9.4	2668	5	US-09-724-676-34549	Sequence 34549, A
c 36	111	9.4	2668	5	US-09-724-676A-34549	Sequence 34549, A
c 37	111	9.4	2768	5	US-09-724-676-34542	Sequence 34542, A
c 38	111	9.4	2768	5	US-09-724-676A-34542	Sequence 34542, A
c 39	111	9.4	3164	5	US-09-724-676-34566	Sequence 34566, A
c 40	111	9.4	3164	5	US-09-724-676A-34566	Sequence 34566, A
c 41	111	9.4	3253	5	US-09-724-676-34525	Sequence 34525, A
c 42	111	9.4	3253	5	US-09-724-676A-34525	Sequence 34525, A
c 43	111	9.4	3264	5	US-09-724-676-34558	Sequence 34558, A
c 44	111	9.4	3264	5	US-09-724-676A-34558	Sequence 34558, A
c 45	111	9.4	3280	5	US-09-724-676-34582	Sequence 34582, A

ALIGNMENTS

RESULT 1
US-09-658-659B-7
; Sequence 7, Application US/09658659B
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: GENE SEQUENCE VARIANCE IN GENES RELATED TO FOLATE METABOLISM HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015001
; CURRENT APPLICATION NUMBER: US/09/658, 659B
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596, 033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357, 743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357, 024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093, 484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 276, 321, 534, 656
; OTHER INFORMATION: n = c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 452, 640
; OTHER INFORMATION: n = a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 492, 625
; OTHER INFORMATION: n = c or a
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 458
; OTHER INFORMATION: nucleotide in position 458 is c, or absent
US-09-658-659B-7

Query Match 100.0%; Score 1187; DB 5; Length 1187;
Best Local Similarity 100.0%; Pred. No. 6.4e-315;
Matches 1187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GATCGCGCACTGCCTCCAGCTGGTGGAGAGACGAGACTGTCTCAAAAAAAAAA 60
|||||

Db 1 GATCGCGCACTGCTCCAGCTGGGTGAGAGAGGAGACTCTGTCTCAAAAAAAAA 60
QY 61 AAAAAAGCCCGAGGCTCAACAAAAACCTCGAAGAGCCCTGCGGTCTTTT 120
Db 61 AAAAAAGCCCGAGGCTCAACAAAAACCTCGAAGAGCCCTGCGGTCTTTT 120
QY 121 TTTTTTTTTTTTTTTTTTGGGACAGTCTTGCTGTGCTGCGCCAGGCTGAGTACATGGT 180
Db 121 TTTTTTTTTTTTTTTTTTGGGACAGTCTTGCTGTGCTGCGCCAGGCTGAGTACATGGT 180
QY 181 CGGATCTGGGTCTACTGACACTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 240
Db 181 CGGATCTGGGTCTACTGACACTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 240
QY 241 TCCCAAGTAGCCACGACCGCCAGCTAATTTTGTATTTTGTAGAGAGCGGGTTTCA 300
Db 241 TCCCAAGTAGCCACGACCGCCAGCTAATTTTGTATTTTGTAGAGAGCGGGTTTCA 300
QY 301 CATGTTGTCAGGCTGCTGTAACCTCTGACCTCAGTGATCCACCGCTCGGCCCCC 360
Db 301 CATGTTGTCAGGCTGCTGTAACCTCTGACCTCAGTGATCCACCGCTCGGCCCCC 360
QY 361 CAAAGTACTAGGATTACAGCGGTGAGCCACCGCTCCAGCGCTTGGCGGTTTTTAATCA 420
Db 361 CAAAGTACTAGGATTACAGCGGTGAGCCACCGCTCCAGCGCTTGGCGGTTTTTAATCA 420
QY 421 AGTAGAAAAGCTGATTAATACCACTTGCTTCGTTGCTGCTGAGAGAGCAAGAAATGG 480
Db 421 AGTAGAAAAGCTGATTAATACCACTTGCTTCGTTGCTGCTGAGAGAGCAAGAAATGG 480
QY 481 AAATGCAATNCNTTATTAGTTGTAGAAACAGATCTCAACACAGTCTTTGTGACAG 540
Db 481 AAATGCAATNCNTTATTAGTTGTAGAAACAGATCTCAACACAGTCTTTGTGACAG 540
QY 541 ACCGAGGAAACGTTGGAACTGTGCTGTAGAGAGGCGGTGACCAAGCG 600
Db 541 ACCGAGGAAACGTTGGAACTGTGCTGTAGAGAGGCGGTGACCAAGCG 600
QY 601 TTCCAAAGGCGGAGTCTTCCCGCCACCGCACTGCTCCAGGTTCCCGGTTTCT 660
Db 601 TTCCAAAGGCGGAGTCTTCCCGCCACCGCACTGCTCCAGGTTCCCGGTTTCT 660
QY 661 AAGACTCTCAGCTGTGCGCTTCCGCTTCTGTCGACACCGCTGCTGCTGCTTTC 720
Db 661 AAGACTCTCAGCTGTGCGCTTCCGCTTCTGTCGACACCGCTGCTGCTGCTTTC 720
QY 721 CCCTGGCGCAGCTCTTAGAGCGGGGCGCGCGAGCCCGCTGCGAGAGAGAGCG 780
Db 721 CCCTGGCGCAGCTCTTAGAGCGGGGCGCGCGAGCCCGCTGCGAGAGAGAGCG 780
QY 781 GAGCGGGAAGCGCGCGGAAAGCGCGGAGGGGTCTGCAACCGCGCACTTGG 840
Db 781 GAGCGGGAAGCGCGCGGAAAGCGCGGAGGGGTCTGCAACCGCGCACTTGG 840
QY 841 CTTGCTCTGCTCCCGCGCGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 841 CTTGCTCTGCTCCCGCGCGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 901 CTTGCTCTGCTCCCGCGCGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 CTTGCTCTGCTCCCGCGCGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 961 TTGCCCCCGCGCGCAGAGAGCGGAGCGCGAGCGGCTGCGAGCTGCGCGCGCG 1020
Db 961 TTGCCCCCGCGCGCAGAGAGCGGAGCGCGAGCGGCTGCGAGCTGCGCGCGCG 1020
QY 1021 TACTGCGGAGATCAACATCTCTGCTGCGGCTGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 TACTGCGGAGATCAACATCTCTGCTGCGGCTGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 ACCGGAACCTGTGCGGTTTTCGATATGAGCGCGCTACAGCTGAGAGGTGAGCGCG 1140
Db 1081 ACCGGAACCTGTGCGGTTTTCGATATGAGCGCGCTACAGCTGAGAGGTGAGCGCG 1140

QY 1141 GGCCCTCGGACGGGTGGCGGAAGAGGAGGAGGCGGCGCTGGGA 1187
Db 1141 GGCCCTCGGACGGGTGGCGGAAGAGGAGGAGGCGGCGCTGGGA 1187
RESULT 2
US-09-658-659B-8
; Sequence 8, Application US/09658659B
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: GENE SEQUENCE VARIANCE IN GENES RELATED
; TITLE OF INVENTION: TO FOLATE METABOLISM HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015001
; CURRENT APPLICATION NUMBER: US/09/658,659B
; CURRENT FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 18597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 701, 13751
; OTHER INFORMATION: n = c or a
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 716, 1293, 2401, 2429, 2618, 3083, 3125, 3635, 4256, 4898,
; LOCATION: 5062, 5167, 11069, 13298, 14479, 14730, 14796, 15344, 15450,
; LOCATION: 15503, 15590, 15840, 16149
; OTHER INFORMATION: n = a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 722, 1379, 1590, 2488, 3212, 5006, 11238, 11422, 11686,
; LOCATION: 12598, 13171, 13645, 13782, 13806, 13813, 14586, 14788,
; LOCATION: 15042, 15546, 15770
; OTHER INFORMATION: n = c or t
; NAME/KEY: misc_feature
; LOCATION: 1322, 1688
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 2594, 11293, 16199, 16203
; OTHER INFORMATION: n = g or t
; NAME/KEY: misc_feature
; LOCATION: 3619-
; OTHER INFORMATION: n = a or t
; NAME/KEY: misc_feature
; LOCATION: 14547
; OTHER INFORMATION: nucleotide in position 14547 is t, or absent
US-09-658-659B-8

Query Match 68.3%; Score 811; DB 5; Length 18597;
Best Local Similarity 99.5%; Pred. No. 4.6e-213;
Matches 1181; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 1 GATCGCGCACTGCTCCAGCTGGGTGAGAGAGGAGACTCTGTCTCAAAAAAAAA 60
Db 78 GATCGCGCACTGCTCCAGCTGGGTGAGAGAGGAGACTCTGTCTCAAAAAAAAA 137

61 AAAAGACCGCCAGGGCTCAAAACAAAACCTCGAAGAACCTCGCGGCTCTTTTTTTT 120
138 AAAAGACCGCCAGGGCTCAAAACAAAACCTCGAAGAACCTCGCGGCTCTTTTTTTT 197
121 TTTTCTTTTCTTTTCTTTTGGACAGTCTTGCTCTGTGCGCCAGGCTGGAGTACAATGGT 180
198 TTTTCTTTTCTTTTCTTTTGGACAGTCTTGCTCTGTGCGCCAGGCTGGAGTACAATGGT 257
181 CGAGCTTTGGCTCACTGCAACCTCTGCTCCCAAGTTCAGCAATCTTCTGCTCAGCC 240
258 CGAGCTTTGGCTCACTGCAACCTCTGCTCCCAAGTTCAGCAATCTTCTGCTCAGCC 317
241 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTATTTAGTAGAGACGGGGTTTACAC 300
318 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTATTTAGTAGAGACGGGGTTTACAC 377
301 CATGTTGTCAGGCTGTCTGACCTCTGCTGAGTATCCACCGGCTCGGCCCCC 360
378 CATGTTGTCAGGCTGTCTGACCTCTGCTGAGTATCCACCGGCTCGGCCCCC 437
361 CAAAGTACTAGGATTAAGGCGGTGAGCCACCGGCTCCAGCGCCCTGGCGGTTTTTAATCA 420
438 CAAAGTACTAGGATTAAGGCGGTGAGCCACCGGCTCCAGCGCCCTGGCGGTTTTTAATCA 497
421 AGTAGAAAGCTGCAATTAACCACTTGTCTTCTGTTGCTTCAAGTGAAGAGAAATGG 480
498 AGTAGAAAGCTGCAATTAACCACTTGTCTTCTGTTGCTTCAAGTGAAGAGAAATGG 556
481 AAATGCAAAATCNCCTTATTAGTGTAGAAACAGATCTCAACAGCAGTTTGTGACAG 540
557 AAATGCAAAATCNCCTTATTAGTGTAGAAACAGATCTCAACAGCAGTTTGTGACAG 616
541 ACCGAGGAAACGTTGGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
617 ACCGAGGAAACGTTGGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
601 TTCCCAAGGCGAGCTCTTCCGCGGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
677 TTCCCAAGGCGAGCTCTTCCGCGGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
661 AGACTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
737 AGACTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
721 CCCTGCGGCACTCTCTAGAGCGGGGCGCGCGGCAACCGGCTGCTGCTGCTGCTGCTGCT 780
797 CCCTGCGGCACTCTCTAGAGCGGGGCGCGCGGCAACCGGCTGCTGCTGCTGCTGCTGCT 856
781 GAGCGGGGACGCGCGCGGGAAGAGCGCGGAGGGGTCTGCTGCTGCTGCTGCTGCTGCTGCT 840
857 GAGCGGGGACGCGCGCGGGAAGAGCGCGGAGGGGTCTGCTGCTGCTGCTGCTGCTGCTGCT 916
841 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
917 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
901 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
977 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
961 TTGCCCCCGCGGACGAGCGGAGCGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
1037 TTGCCCCCGCGGACGAGCGGAGCGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1096
1021 TACTGGGCGAGATCCACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
1097 TACTGGGCGAGATCCACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156
1081 ACAGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
1157 ACAGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1216
1141 GGCCTCTGCGGACGGGTGCGGGAAGAGGAGGCGCGGCTGGGGA 1187

Db 1217 GCGCCCTCGCGGACGGGTGCGGGAAGGAGGAGCGCGGCTGGGA 1263
RESULT 3
US-09-724-676-34523/c
; Sequence 34523, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34523
; LENGTH: 1942
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-34523
Query Match 25.4%; Score 301; DB 5; Length 1942;
Best Local Similarity 100.0%; Pred. No. 4e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 833 CCACCTTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 892
Db 1929 CCACCTTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1870
Qy 893 TTGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 952
Db 1869 TTGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1810
Qy 953 GCGCGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1012
Db 1809 GCGCGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1750
Qy 1013 AGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1072
Db 1749 AGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1690
Qy 1073 GCAAGGCGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1132
Db 1689 GCAAGGCGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1630
Qy 1133 A 1133
Db 1629 A 1629
RESULT 4
US-09-724-676A-34523/c
; Sequence 34523, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34523
; LENGTH: 1942
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-34523
Query Match 25.4%; Score 301; DB 5; Length 1942;
Best Local Similarity 100.0%; Pred. No. 4e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 833 CCACCTTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 892

```
Db 1929 CCACTTGGCTGCTCGCTCCGCGCGCCACATTGGGCTGCTCCGTCGCCCGCGCCAC 1870
QY 893 TTGCGCTGCTCGCTCCGCGCGCCCGCGCCGCAATGCTGTCGCGCTCGAGCTGCCGC 952
Db 1869 TTGCGCTGCTCGCTCCGCGCGCCCGCGCCGCAATGCTGTCGCGCTCGAGCTGCCGC 1810
QY 953 GCGGCGCTTGGCCCGCGCGCCGCGCAGAGAGCGGAGCGCGGCTGTCGCGCGCAGCGGG 1012
Db 1809 GCGGCGCTTGGCCCGCGCGCCGCGCAGAGAGCGGAGCGCGGCTGTCGCGCGCAGCGGG 1750
QY 1013 AGCTGCGATGCTGGGCGCAGATCCAAACATCTCCGCTCCGCTCGCGCTCAGGAGGACGACC 1072
Db 1749 AGCTGCGATGCTGGGCGCAGATCCAAACATCTCCGCTCCGCTCGCGCTCAGGAGGACGACC 1690
QY 1073 GCACGGGACCGGACCGCTGTCGCTGATTTGCGCATGCGGCGCGCTACAGCTTGAGAGTG 1132
Db 1689 GCACGGGACCGGACCGCTGTCGCTGATTTGCGCATGCGGCGCGCTACAGCTTGAGAGTG 1630
QY 1133 A 1133
Db 1629 A 1629

RESULT 5
US-09-724-676-34515/c
; Sequence 34515, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34515
; LENGTH: 2042
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-34515

Query Match 25.4%; Score 301; DB 5; Length 2042;
Best Local Similarity 100.0%; Pred. No. 3.9e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 833 CCACTTGGCTGCTCGCTCCGCGCGCCACATTGGGCTGCTCCGTCGCCCGCGCCAC 892
Db 2029 CCACTTGGCTGCTCGCTCCGCGCGCCACATTGGGCTGCTCCGTCGCCCGCGCCAC 1970
QY 893 TTGCGCTGCTCGCTCCGCGCGCCCGCGCCGCAATGCTGTCGCGCTCGAGCTGCCGC 952
Db 1969 TTGCGCTGCTCGCTCCGCGCGCCCGCGCCGCAATGCTGTCGCGCTCGAGCTGCCGC 1910
QY 953 GCGGCGCTTGGCCCGCGCGCCGCGCAGAGAGCGGAGCGCGGCTGTCGCGCGCAGCGGG 1012
Db 1909 GCGGCGCTTGGCCCGCGCGCCGCGCAGAGAGCGGAGCGCGGCTGTCGCGCGCAGCGGG 1850
QY 1013 AGCTGCGATGCTGGGCGCAGATCCAAACATCTCCGCTCCGCTCGCGCTCAGGAGGACGACC 1072
Db 1849 AGCTGCGATGCTGGGCGCAGATCCAAACATCTCCGCTCCGCTCGCGCTCAGGAGGACGACC 1790
QY 1073 GCACGGGACCGGACCGCTGTCGCTGATTTGCGCATGCGGCGCGCTACAGCTTGAGAGTG 1132
Db 1789 GCACGGGACCGGACCGCTGTCGCTGATTTGCGCATGCGGCGCGCTACAGCTTGAGAGTG 1730
QY 1133 A 1133
Db 1729 A 1729

Query Match 25.4%; Score 301; DB 5; Length 2042;
Best Local Similarity 100.0%; Pred. No. 3.9e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 833 CCACTTGGCTGCTCGCTCCGCGCGCCACATTGGGCTGCTCCGTCGCCCGCGCCAC 892
Db 2029 CCACTTGGCTGCTCGCTCCGCGCGCCACATTGGGCTGCTCCGTCGCCCGCGCCAC 1970
QY 893 TTGCGCTGCTCGCTCCGCGCGCCCGCGCCGCAATGCTGTCGCGCTCGAGCTGCCGC 952
Db 1969 TTGCGCTGCTCGCTCCGCGCGCCCGCGCCGCAATGCTGTCGCGCTCGAGCTGCCGC 1910
QY 953 GCGGCGCTTGGCCCGCGCGCCGCGCAGAGAGCGGAGCGCGGCTGTCGCGCGCAGCGGG 1012
Db 1909 GCGGCGCTTGGCCCGCGCGCCGCGCAGAGAGCGGAGCGCGGCTGTCGCGCGCAGCGGG 1850
QY 1013 AGCTGCGATGCTGGGCGCAGATCCAAACATCTCCGCTCCGCTCGCGCTCAGGAGGACGACC 1072
Db 1849 AGCTGCGATGCTGGGCGCAGATCCAAACATCTCCGCTCCGCTCGCGCTCAGGAGGACGACC 1790
QY 1073 GCACGGGACCGGACCGCTGTCGCTGATTTGCGCATGCGGCGCGCTACAGCTTGAGAGTG 1132
Db 1789 GCACGGGACCGGACCGCTGTCGCTGATTTGCGCATGCGGCGCGCTACAGCTTGAGAGTG 1730
QY 1133 A 1133
Db 1729 A 1729

RESULT 6
US-09-724-676A-34515/c
; Sequence 34515, Application US/09724676A
; GENERAL INFORMATION:
```

```
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34515
; LENGTH: 2042
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-34515

Query Match 25.4%; Score 301; DB 5; Length 2042;
Best Local Similarity 100.0%; Pred. No. 3.9e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 833 CCACTTGGCTGCTCGCTCCGCGCGCCACATTGGGCTGCTCCGTCGCCCGCGCCAC 892
Db 2029 CCACTTGGCTGCTCGCTCCGCGCGCCACATTGGGCTGCTCCGTCGCCCGCGCCAC 1970
QY 893 TTGCGCTGCTCGCTCCGCGCGCCCGCGCCGCAATGCTGTCGCGCTCGAGCTGCCGC 952
Db 1969 TTGCGCTGCTCGCTCCGCGCGCCCGCGCCGCAATGCTGTCGCGCTCGAGCTGCCGC 1910
QY 953 GCGGCGCTTGGCCCGCGCGCCGCGCAGAGAGCGGAGCGCGGCTGTCGCGCGCAGCGGG 1012
Db 1909 GCGGCGCTTGGCCCGCGCGCCGCGCAGAGAGCGGAGCGCGGCTGTCGCGCGCAGCGGG 1850
QY 1013 AGCTGCGATGCTGGGCGCAGATCCAAACATCTCCGCTCCGCTCGCGCTCAGGAGGACGACC 1072
Db 1849 AGCTGCGATGCTGGGCGCAGATCCAAACATCTCCGCTCCGCTCGCGCTCAGGAGGACGACC 1790
QY 1073 GCACGGGACCGGACCGCTGTCGCTGATTTGCGCATGCGGCGCGCTACAGCTTGAGAGTG 1132
Db 1789 GCACGGGACCGGACCGCTGTCGCTGATTTGCGCATGCGGCGCGCTACAGCTTGAGAGTG 1730
QY 1133 A 1133
Db 1729 A 1729

RESULT 7
US-09-724-676-34597/c
; Sequence 34597, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34597
; LENGTH: 2391
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-34597

Query Match 25.4%; Score 301; DB 5; Length 2391;
Best Local Similarity 100.0%; Pred. No. 3.7e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 833 CCACTTGGCTGCTCGCTCCGCGCGCCACATTGGGCTGCTCCGTCGCCCGCGCCAC 892
Db 2378 CCACTTGGCTGCTCGCTCCGCGCGCCACATTGGGCTGCTCCGTCGCCCGCGCCAC 2319
QY 893 TTGCGCTGCTCGCTCCGCGCGCCCGCGCCGCAATGCTGTCGCGCTCGAGCTGCCGC 952
Db 2318 TTGCGCTGCTCGCTCCGCGCGCCCGCGCCGCAATGCTGTCGCGCTCGAGCTGCCGC 2259
QY 953 GCGGCGCTTGGCCCGCGCGCCGCGCAGAGAGCGGAGCGCGGCTGTCGCGCGCAGCGGG 1012
```

b 2258 GCGGCGCTTCCCGCGCGCACAGAGCGGGACGCGCGCTCCCGCGCACGCGG 2199
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (41)..(41)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-34535

Query Match 25.4%; Score 301; DB 5; Length 2411;
Best Local Similarity 100.0%; Pred. No. 3.7e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1013 AGCTCAGTACTGGGCGAGATCCACATCTCCGTCGGCGTCAGAGAGCGACC 1072
b 2198 AGCTCAGTACTGGGCGAGATCCACATCTCCGTCGGCGTCAGAGAGCGACC 2139
Y 1073 GCACGGGACCGGACCGCTCGGTATTCGGCATCGAGCGCGCTACAGCTGAGAGGTG 1132
b 2138 GCACGGGACCGGACCGCTCGGTATTCGGCATCGAGCGCGCTACAGCTGAGAGGTG 2079
Y 1133 A 1133
b 2078 A 2078

RESULT 8
US-09-724-676A-34597/c
Sequence 34597, Application US/09724676A
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676A
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 34597
LENGTH: 2391
TYPE: DNA
ORGANISM: Homo sapiens
US-09-724-676A-34597

Query Match 25.4%; Score 301; DB 5; Length 2391;
Best Local Similarity 100.0%; Pred. No. 3.7e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 833 CCACCTTGGCTGCTCCGTCGCGCGCGCACATTCGCTCGCTCCCGCGCGCAC 892
b 2378 CCACCTTGGCTGCTCCGTCGCGCGCGCACATTCGCTCGCTCCCGCGCGCAC 2319
Y 893 TTGCGCTGCTCCGTCGCGCGCGCACATTCGCTCGCTCCCGCGCGCAC 952
b 2318 TTGCGCTGCTCCGTCGCGCGCGCACATTCGCTCGCTCCCGCGCGCAC 2259
Y 953 GCGGCGCTTGGCGCGCGCGCACAGGAGCGGACGCGCGCTCCGCGCACGCGG 1012
b 2258 GCGGCGCTTGGCGCGCGCGCACAGGAGCGGACGCGCGCTCCGCGCACGCGG 2199
Y 1013 AGCTCAGTACTGGGCGAGATCCACATCTCCGTCGGCGCTCAGGAGGACGACC 1072
b 2198 AGCTCAGTACTGGGCGAGATCCACATCTCCGTCGGCGCTCAGGAGGACGACC 2139
Y 1073 GCACGGGACCGGACCGCTCGGTATTCGGCATCGAGCGCGCTACAGCTGAGAGGTG 1132
b 2138 GCACGGGACCGGACCGCTCGGTATTCGGCATCGAGCGCGCTACAGCTGAGAGGTG 2079
Y 1133 A 1133
b 2078 A 2078

RESULT 9
US-09-724-676-34535/c
Sequence 34535, Application US/09724676
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 34535

Y 833 CCACCTTGGCTGCTCCGTCGCGCGCGCACATTCGCTCGCTCCCGCGCGCAC 892
b 2398 CCACCTTGGCTGCTCCGTCGCGCGCGCACATTCGCTCGCTCCCGCGCGCAC 2339
Y 893 TTGCGCTGCTCCGTCGCGCGCGCACATTCGCTCGCTCCCGCGCGCAC 952
b 2338 TTGCGCTGCTCCGTCGCGCGCGCACATTCGCTCGCTCCCGCGCGCAC 2279
Y 953 GCGGCGCTTGGCGCGCGCGCACAGGAGCGGACGCGCGCTCCGCGCACGCGG 1012
b 2278 GCGGCGCTTGGCGCGCGCGCACAGGAGCGGACGCGCGCTCCGCGCACGCGG 2219
Y 1013 AGCTCAGTACTGGGCGAGATCCACATCTCCGTCGGCGCTCAGGAGGACGACC 1072
b 2218 AGCTCAGTACTGGGCGAGATCCACATCTCCGTCGGCGCTCAGGAGGACGACC 2159
Y 1073 GCACGGGACCGGACCGCTCGGTATTCGGCATCGAGCGCGCTACAGCTGAGAGGTG 1132
b 2158 GCACGGGACCGGACCGCTCGGTATTCGGCATCGAGCGCGCTACAGCTGAGAGGTG 2099
Y 1133 A 1133
b 2098 A 2098

RESULT 10
US-09-724-676A-34535/c
Sequence 34535, Application US/09724676A
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676A
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 34535
LENGTH: 2411
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (41)..(41)
OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-34535

Query Match 25.4%; Score 301; DB 5; Length 2411;
Best Local Similarity 100.0%; Pred. No. 3.7e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 833 CCACCTTGGCTGCTCCGTCGCGCGCGCACATTCGCTCGCTCCCGCGCGCAC 892
b 2398 CCACCTTGGCTGCTCCGTCGCGCGCGCACATTCGCTCGCTCCCGCGCGCAC 2339
Y 893 TTGCGCTGCTCCGTCGCGCGCGCACATTCGCTCGCTCCCGCGCGCAC 952
b 2338 TTGCGCTGCTCCGTCGCGCGCGCACATTCGCTCGCTCCCGCGCGCAC 2279
Y 953 GCGGCGCTTGGCGCGCGCGCACAGGAGCGGACGCGCGCTCCGCGCACGCGG 1012

```
Db 2278 GCCGGCCCTTGGCCCCCGCCGACAGAGAGCGGAGCGCGGCTCCGCGCCGACCGGGG 2219
Qy 1013 AGCTGCAGTACTGGGGCAGATCCAAACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 1072
Db 2218 AGCTGCAGTACTGGGGCAGATCCAAACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 2159
Qy 1073 GCACGGGACCGGACCCCTGTCGGTATTTCGGCATGCAAGGCGGCTCAGAGCTGAGAGTG 1132
Db 2158 GCACGGGACCGGACCCCTGTCGGTATTTCGGCATGCAAGGCGGCTCAGAGCTGAGAGTG 2099
Qy 1133 A 1133
Db 2098 A 2098

RESULT 11
US-09-724-676-34589/c
; Sequence 34589, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34589
; LENGTH: 2491
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-34589

Query Match      25.4%; Score 301; DB 5; Length 2491;
Best Local Similarity 100.0%; Pred. No. 3.6e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 833 CCACCTTGGCTTGCCTCCGTCCTCCGCGCCGACATCCACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 892
Db 2478 CCACCTTGGCTTGCCTCCGTCCTCCGCGCCGACATCCACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 2419
Qy 893 TTCGCTTGCCTTCCGTCCTCCGCGCCGACATCCACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 952
Db 2418 TTCGCTTGCCTTCCGTCCTCCGCGCCGACATCCACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 2359
Qy 953 GCCGGCCCTTGGCCCCCGCCGACAGAGAGCGGAGCGCGCTCCGCGCCGACCGGGG 1012
Db 2358 GCCGGCCCTTGGCCCCCGCCGACAGAGAGCGGAGCGCGCTCCGCGCCGACCGGGG 2299
Qy 1013 AGCTGCAGTACTGGGGCAGATCCAAACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 1072
Db 2298 AGCTGCAGTACTGGGGCAGATCCAAACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 2239
Qy 1073 GCACGGGACCGGACCCCTGTCGGTATTTCGGCATGCAAGGCGGCTCAGAGCTGAGAGTG 1132
Db 2238 GCACGGGACCGGACCCCTGTCGGTATTTCGGCATGCAAGGCGGCTCAGAGCTGAGAGTG 2179
Qy 1133 A 1133
Db 2178 A 2178

RESULT 13
US-09-724-676-34548/c
; Sequence 34548, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34548
; LENGTH: 2527
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (41)..(41)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-34548

Query Match      25.4%; Score 301; DB 5; Length 2527;
Best Local Similarity 100.0%; Pred. No. 3.6e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 833 CCACCTTGGCTTGCCTCCGTCCTCCGCGCCGACATCCACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 892
Db 2514 CCACCTTGGCTTGCCTCCGTCCTCCGCGCCGACATCCACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 2455
Qy 893 TTCGCTTGCCTTCCGTCCTCCGCGCCGACATCCACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 952
Db 2454 TTCGCTTGCCTTCCGTCCTCCGCGCCGACATCCACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 2395
Qy 953 GCCGGCCCTTGGCCCCCGCCGACAGAGAGCGGAGCGCGCTCCGCGCCGACCGGGG 1012
Db 2394 GCCGGCCCTTGGCCCCCGCCGACAGAGAGCGGAGCGCGCTCCGCGCCGACCGGGG 2335
Qy 1013 AGCTGCAGTACTGGGGCAGATCCAAACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 1072
Db 2178 AGCTGCAGTACTGGGGCAGATCCAAACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 2179
```

```
Db 2278 GCCGGCCCTTGGCCCCCGCCGACAGAGAGCGGAGCGCGGCTCCGCGCCGACCGGGG 2219
Qy 1013 AGCTGCAGTACTGGGGCAGATCCAAACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 1072
Db 2218 AGCTGCAGTACTGGGGCAGATCCAAACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 2159
Qy 1073 GCACGGGACCGGACCCCTGTCGGTATTTCGGCATGCAAGGCGGCTCAGAGCTGAGAGTG 1132
Db 2158 GCACGGGACCGGACCCCTGTCGGTATTTCGGCATGCAAGGCGGCTCAGAGCTGAGAGTG 2099
Qy 1133 A 1133
Db 2098 A 2098

RESULT 11
US-09-724-676-34589/c
; Sequence 34589, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34589
; LENGTH: 2491
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-34589

Query Match      25.4%; Score 301; DB 5; Length 2491;
Best Local Similarity 100.0%; Pred. No. 3.6e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 833 CCACCTTGGCTTGCCTCCGTCCTCCGCGCCGACATCCACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 892
Db 2478 CCACCTTGGCTTGCCTCCGTCCTCCGCGCCGACATCCACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 2419
Qy 893 TTCGCTTGCCTTCCGTCCTCCGCGCCGACATCCACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 952
Db 2418 TTCGCTTGCCTTCCGTCCTCCGCGCCGACATCCACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 2359
Qy 953 GCCGGCCCTTGGCCCCCGCCGACAGAGAGCGGAGCGCGCTCCGCGCCGACCGGGG 1012
Db 2358 GCCGGCCCTTGGCCCCCGCCGACAGAGAGCGGAGCGCGCTCCGCGCCGACCGGGG 2299
Qy 1013 AGCTGCAGTACTGGGGCAGATCCAAACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 1072
Db 2298 AGCTGCAGTACTGGGGCAGATCCAAACATCTCTCGCTGCGGCGCTCAGGAAGACGACC 2239
Qy 1073 GCACGGGACCGGACCCCTGTCGGTATTTCGGCATGCAAGGCGGCTCAGAGCTGAGAGTG 1132
Db 2238 GCACGGGACCGGACCCCTGTCGGTATTTCGGCATGCAAGGCGGCTCAGAGCTGAGAGTG 2179
Qy 1133 A 1133
Db 2178 A 2178

RESULT 12
US-09-724-676A-34589/c
; Sequence 34589, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34589
```



```
Query Match      25.4%; Score 301; DB 5; Length 3023;
Best Local Similarity 100.0%; Pred. No. 3.4e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	833	CAACTTGGCGTGGCTCCGTCGCCGCGGCACTTTGGCTGCTCGTCCGCGCGCGCAC	892
Db	3010	CCACTTGGCGTGGCTCCGTCGCCGCGGCACTTTGGCTGCTCGTCCGCGCGCGCAC	2951
QY	893	TTTCGCTGCTCGGTCCCGCCGCCGCGCATGCTGTGCGCGGCTTCGAGACTGCCGC	952
Db	2950	TTTCGCTGCTCGGTCCCGCCGCCGCGCATGCTGTGCGCGGCTTCGAGACTGCCGC	2891
QY	953	GCGGGCCCTTGGCCCCCGCCGCGCACAGAGCGGAGACCCGAGCGCGCTCCGCCGACACGGG	1012
Db	2890	GCGGGCCCTTGGCCCCCGCCGCGCACAGAGCGGAGACCCGAGCGCGCTCCGCCGACACGGG	2831
QY	1013	AGCTGCAGTACCTGGGGCGATCCAAACATCTCTCGCTTGGCGGTCAGGAAGACGACC	1072
Db	2830	AGCTGCAGTACCTGGGGCGATCCAAACATCTCTCGCTTGGCGGTCAGGAAGACGACC	2771
QY	1073	GCACGGGCAACCGCACCCCTGTTCGGATTTCGGCATTCAGACGGCGCTACAGCTTGAGAGTG	1132
Db	2770	GCAACGGGCAACCGCACCCCTGTTCGGATTTCGGCATTCAGACGGCGCTACAGCTTGAGAGTG	2711
QY	1133	A 1133	
Db	2710	A 2710	

Search completed: November 29, 2002, 05:31:14
Job time : 701.042 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2002, 00:35:25 ; Search time 2462.49 Seconds
(without alignments)
12119.489 Million cell updates/sec

Title: US-09-963-333-7
Perfect score: 1187
Sequence: 1 gatcgccactgactcca.....gagggaggcgcgctgggga 1187

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 24791104 seqs, 12571243825 residues

Word size : 15

Total number of hits satisfying chosen parameters: 2685365

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending Patents NA Main:*

- 1: /cgn2_6/ptodata/2/pna/pctus COMB.seq:*
- 2: /cgn2_6/ptodata/2/pna/us06 COMB.seq:*
- 3: /cgn2_6/ptodata/2/pna/us07 COMB.seq:*
- 4: /cgn2_6/ptodata/2/pna/us08 COMB.seq:*
- 5: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 6: /cgn2_6/ptodata/2/pna/us08 COMB.seq:*
- 7: /cgn2_6/ptodata/2/pna/us08 COMB.seq:*
- 8: /cgn2_6/ptodata/2/pna/us08 COMB.seq:*
- 9: /cgn2_6/ptodata/2/pna/us08 COMB.seq:*
- 10: /cgn2_6/ptodata/2/pna/us08 COMB.seq:*
- 11: /cgn2_6/ptodata/2/pna/us08 COMB.seq:*
- 12: /cgn2_6/ptodata/2/pna/us08 COMB.seq:*
- 13: /cgn2_6/ptodata/2/pna/us08 COMB.seq:*
- 14: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 15: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 16: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 17: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 18: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 19: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 20: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 21: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 22: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 23: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 24: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 25: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 26: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 27: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 28: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 29: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 30: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 31: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 32: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 33: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 34: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 35: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 36: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 37: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 38: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 39: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 40: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 41: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 42: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*
- 43: /cgn2_6/ptodata/2/pna/us09 COMB.seq:*

- 44: /cgn2_6/ptodata/2/pna/us6000 COMB.seq:*
- 45: /cgn2_6/ptodata/2/pna/us6001 COMB.seq:*
- 46: /cgn2_6/ptodata/2/pna/us6002 COMB.seq:*
- 47: /cgn2_6/ptodata/2/pna/us6003 COMB.seq:*
- 48: /cgn2_6/ptodata/2/pna/us6004 COMB.seq:*
- 49: /cgn2_6/ptodata/2/pna/us6005 COMB.seq:*
- 50: /cgn2_6/ptodata/2/pna/us6006 COMB.seq:*
- 51: /cgn2_6/ptodata/2/pna/us6007 COMB.seq:*
- 52: /cgn2_6/ptodata/2/pna/us6008 COMB.seq:*
- 53: /cgn2_6/ptodata/2/pna/us6009 COMB.seq:*
- 54: /cgn2_6/ptodata/2/pna/us6010 COMB.seq:*
- 55: /cgn2_6/ptodata/2/pna/us6011 COMB.seq:*
- 56: /cgn2_6/ptodata/2/pna/us6012 COMB.seq:*
- 57: /cgn2_6/ptodata/2/pna/us6013 COMB.seq:*
- 58: /cgn2_6/ptodata/2/pna/us6014 COMB.seq:*
- 59: /cgn2_6/ptodata/2/pna/us6015 COMB.seq:*
- 60: /cgn2_6/ptodata/2/pna/us6016 COMB.seq:*
- 61: /cgn2_6/ptodata/2/pna/us6017 COMB.seq:*
- 62: /cgn2_6/ptodata/2/pna/us6018 COMB.seq:*
- 63: /cgn2_6/ptodata/2/pna/us6019 COMB.seq:*
- 64: /cgn2_6/ptodata/2/pna/us6020 COMB.seq:*
- 65: /cgn2_6/ptodata/2/pna/us6021 COMB.seq:*
- 66: /cgn2_6/ptodata/2/pna/us6022 COMB.seq:*
- 67: /cgn2_6/ptodata/2/pna/us6023 COMB.seq:*
- 68: /cgn2_6/ptodata/2/pna/us6024 COMB.seq:*
- 69: /cgn2_6/ptodata/2/pna/us6025 COMB.seq:*
- 70: /cgn2_6/ptodata/2/pna/us6026 COMB.seq:*
- 71: /cgn2_6/ptodata/2/pna/us6027 COMB.seq:*
- 72: /cgn2_6/ptodata/2/pna/us6028 COMB.seq:*
- 73: /cgn2_6/ptodata/2/pna/us6029 COMB.seq:*
- 74: /cgn2_6/ptodata/2/pna/us6030 COMB.seq:*
- 75: /cgn2_6/ptodata/2/pna/us6031 COMB.seq:*
- 76: /cgn2_6/ptodata/2/pna/us6032 COMB.seq:*
- 77: /cgn2_6/ptodata/2/pna/us6033 COMB.seq:*
- 78: /cgn2_6/ptodata/2/pna/us6034 COMB.seq:*
- 79: /cgn2_6/ptodata/2/pna/us6035 COMB.seq:*
- 80: /cgn2_6/ptodata/2/pna/us6036 COMB.seq:*
- 81: /cgn2_6/ptodata/2/pna/us6037 COMB.seq:*
- 82: /cgn2_6/ptodata/2/pna/us6038 COMB.seq:*
- 83: /cgn2_6/ptodata/2/pna/us6039 COMB.seq:*
- 84: /cgn2_6/ptodata/2/pna/us6040 COMB.seq:*
- 85: /cgn2_6/ptodata/2/pna/us6041 COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1187	100.0	1187	25	US-09-658-659-7 Sequence 7, Appli
2	1187	100.0	1187	36	US-09-962-665-7 Sequence 7, Appli
3	1187	100.0	1187	36	US-09-962-677-7 Sequence 7, Appli
4	1187	100.0	1187	36	US-09-963-333-7 Sequence 7, Appli
5	811	68.3	18597	25	US-09-658-659-8 Sequence 8, Appli
6	811	68.3	18597	36	US-09-962-665-8 Sequence 8, Appli
7	811	68.3	18597	36	US-09-962-677-8 Sequence 8, Appli
8	811	68.3	18597	36	US-09-963-333-8 Sequence 8, Appli
9	658	55.4	18596	24	US-09-577-266-11 Sequence 11, Appl
10	658	55.4	18596	24	US-09-631-275-140 Sequence 140, App
11	658	55.4	18596	33	US-09-880-107-1590 Sequence 1590, App
12	658	55.4	18596	36	US-09-934-531-124 Sequence 124, App
13	658	55.4	18596	36	US-09-954-531-348 Sequence 348, App
14	658	55.4	18596	36	US-09-967-768A-119 Sequence 119, App
15	297	25.0	1536	1	PCT-US02-18947-556 Sequence 556, App
16	297	25.0	1536	1	PCT-US99-12080-1 Sequence 1, Appli
17	297	25.0	1536	18	US-09-442-384A-792 Sequence 792, App
18	297	25.0	1536	18	US-09-442-384B-792 Sequence 792, App
19	297	25.0	1536	24	US-09-631-275-141 Sequence 141, App
20	297	25.0	1536	25	US-09-658-659-6 Sequence 6, Appli
21	297	25.0	1536	33	US-09-877-178-11 Sequence 11, Appli

Db 1081 ACCGGACCCCTGTCGGTATTCGGCATGCGGCGGCTACAGCTGAGAGTGCACGCGG 1140
QY 1141 GCGCCCTGCGGGACGGGTGCGGGAAGAGAGGAGGCGCGCTGGGA 1187
Db 1141 GCGCCCTGCGGGACGGGTGCGGGAAGAGAGGAGGCGCGCTGGGA 1187

RESULT 2

US-09-962-665-7
; Sequence 7, Application US/09962665
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: POLY(POLYGLUTAMATE SYNTHETASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015004
; CURRENT APPLICATION NUMBER: US/09/962,665
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 276, 321, 534, 656
; OTHER INFORMATION: n = c or t
; NAME/KEY: misc feature
; LOCATION: 452-640
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc feature
; LOCATION: 492, 625
; OTHER INFORMATION: n = c or a
; NAME/KEY: misc feature
; LOCATION: 458
; OTHER INFORMATION: nucleotide in position 458 is c, or absent
US-09-962-665-7

Query Match 100.0%; Score 1187; DB 36; Length 1187;
Best Local Similarity 100.0%; Pred. No. 6.6e-139;
Matches 1187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GATCGCGCCACTGCACTCCAGCTGGGTGAGAGCGAGCTCTCTCTCAAAAAA 60
Db 1 GATCGCGCCACTGCACTCCAGCTGGGTGAGAGCGAGCTCTCTCTCAAAAAA 60
QY 61 AAAAAAGCCGCGAGGCTCAACAAAAAAGCCCTGGGAGGCTCTCTCTCTCTTTT 120
Db 61 AAAAAAGCCGCGAGGCTCAACAAAAAAGCCCTGGGAGGCTCTCTCTCTCTTTT 120
QY 121 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 180
Db 121 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 180
QY 181 CGGATCTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTGCTCAGCC 240
Db 181 CGGATCTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTGCTCAGCC 240
QY 241 TCCCAAGTAGCCACCAAGCCAGCTTAATTTTGTANTTTTAGTAGAGCGGGGTTTAC 300
Db 241 TCCCAAGTAGCCACCAAGCCAGCTTAATTTTGTANTTTTAGTAGAGCGGGGTTTAC 300

QY 301 CATGTTGTCCAGGTGGTCTNGAACTCTGACCTCAGGTGATCCACCCGCTCGGCCCCC 360
Db 301 CATGTTGTCCAGGTGGTCTNGAACTCTGACCTCAGGTGATCCACCCGCTCGGCCCCC 360
QY 361 CAAAGTACTAGGATTAAGCGGTGAGCCACCGCTTCAGCGCCCTGCGCGTCTTTAAATCA 420
Db 361 CAAAGTACTAGGATTAAGCGGTGAGCCACCGCTTCAGCGCCCTGCGCGTCTTTAAATCA 420
QY 421 AGTAGAAAAGCTGCATTTATACCACTTGTCTTNGTTTGTCTTCTAGTGAGAACGAAATGG 480
Db 421 AGTAGAAAAGCTGCATTTATACCACTTGTCTTNGTTTGTCTTCTAGTGAGAACGAAATGG 480
QY 481 AAATGCAATCNCCTTATTAGTCTAGGAAACAGATCTCAACAGCAGTCTTTGTNGACAG 540
Db 481 AAATGCAATCNCCTTATTAGTCTAGGAAACAGATCTCAACAGCAGTCTTTGTNGACAG 540
QY 541 ACCGAGGAAAACGTTGGGAACTGTGTCTGTAGGAAAGGCGCGTGCACCAAGG 600
Db 541 ACCGAGGAAAACGTTGGGAACTGTGTCTGTAGGAAAGGCGCGTGCACCAAGG 600
QY 601 TTCCAAAAGGCGCAGTCTTCCNGGCAACCGCAGCTGCTTCCAGGTTCCCGGTTTCT 660
Db 601 TTCCAAAAGGCGCAGTCTTCCNGGCAACCGCAGCTGCTTCCAGGTTCCCGGTTTCT 660
QY 661 AAGACTCTCAGTGTGGCCCTTGGGCTTCTGTGCCACACCCCTGCTCTCTCGTTTC 720
Db 661 AAGACTCTCAGTGTGGCCCTTGGGCTTCTGTGCCACACCCCTGCTCTCTCGTTTC 720
QY 721 CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGCAGCTCCAGGTTCCCGGTTTCT 780
Db 721 CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGCAGCTCCAGGTTCCCGGTTTCT 780
QY 781 GAGCGCGGACCGCGCGGGAAGGCGCGGAGGGGTCTGCTCACCGGCGCACTTGG 840
Db 781 GAGCGCGGACCGCGCGGGAAGGCGCGGAGGGGTCTGCTCACCGGCGCACTTGG 840
QY 841 CCTGCTCTCGCTCCCGCGCGGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 841 CCTGCTCTCGCTCCCGCGCGGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 901 CCTGCTCTCGCTCCCGCGCGGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 CCTGCTCTCGCTCCCGCGCGGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 961 TTGCCCCCGCGCGCAGGAGGCGCGCGCGCTGCGAGCTGCGCGCGCGCGCC 1020
Db 961 TTGCCCCCGCGCGCAGGAGGCGCGCGCGCTGCGAGCTGCGCGCGCGCGCC 1020
QY 1021 TACTTGGGCGAGATCCACACATCTCTGCTGCGGCTCAGGAGGACGCGCAGCGGC 1080
Db 1021 TACTTGGGCGAGATCCACACATCTCTGCTGCGGCTCAGGAGGACGCGCAGCGGC 1080
QY 1081 ACCGCGACCTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 1081 ACCGCGACCTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
QY 1141 GCGCCCTGCGGAGCGGCTGCGGGAAGGAGGAGGCGCGCTGGGGA 1187
Db 1141 GCGCCCTGCGGAGCGGCTGCGGGAAGGAGGAGGCGCGCTGGGGA 1187

RESULT 3

US-09-962-677-7
; Sequence 7, Application US/09962677
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
; TITLE OF INVENTION: THE TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015003
; CURRENT APPLICATION NUMBER: US/09/962,677
; CURRENT FILING DATE: 2001-09-24

QY	1	GATCGCCCACTGCATCTCCAGCCTGGGGTGGAGAGCGAGACTCTGTCTCAAAAAA	60
Db	1		
QY	61	AAAAAGACCGCCAGGGCTCAAAACAAAAAACCCTGGAAAGCCCTGGCGGTCTTTTTTTTT	120
Db	61		
QY	121	TTTTTTTTTTTTTTTTTTTTTTGGGAAGCTTCTCTCTGCGCCAGGCTGGAGTCAATGTT	180
Db	121		
QY	181	CGGATCTTGGCTCATCTGAACTCTGCTCTCCAGGTTTCAAGCAATTTCTTGCCTCAGCC	240
Db	181		
QY	241	TCCCAAGTAGCCACCAAGCCGAGCTAATTTTTGTATNTTTAGTAGAGACGGGGTTTCAC	300
Db	241		
QY	301	CATGTTGCCAGGCTGGTCTNGAACTCTCGACTCAGGTGATCCACCGCCTCGCGCCCC	360
Db	301		
QY	361	CAAGTACTAGGATTACAGCGGTGAGCACCGGCTCCAGCGCCCTGGCGGTTTTTAATCA	420
Db	361		
QY	421	AGTAGAAAAGCTGCATTATACCATCTTGCTTCTGCTTCCAGTGAGAACGAAATGG	480
Db	421		
QY	481	AAATGCAAAATCNCATTATAGTTGTAGGAAACAGATCTAAACACAGTTTTCTTGACAG	540
Db	481		
QY	541	ACCGCAGGAAAACTGGGAACTGTGCTGCTGGCTTGGTAGAGAGCGCGGTCTGACACAGCG	600

```

; SEQ ID NO 1
; LENGTH: 1187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 276..321, 534..656

```

OTHER INFORMATION: n = c or t
NAME/KEY: misc_feature
LOCATION: 452, 640
OTHER INFORMATION: n = a or g
NAME/KEY: misc_feature
LOCATION: 492, 625
OTHER INFORMATION: n = c or a
NAME/KEY: misc_feature
LOCATION: 458
OTHER INFORMATION: nucleotide in position 458 is c, or absent

US-09-963-333-7
Query Match 100.0%; Score 1187; DB 36; Length 1187;
Best Local Similarity 100.0%; Pred. No. 6.6e-139;
Matches 1187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GATCGCCACTGCTCCTCAGCTGGGTGAGAGAGCAGACTCTGCTCAAAAAA 60
1 GATCGCCACTGCTCCTCAGCTGGGTGAGAGAGCAGACTCTGCTCAAAAAA 60
61 AAAAAGACCCGAGGCTCAAAACAAAACCTCGAAAAGCCCTGCGGTCTTTTTT 120
61 AAAAAGACCCGAGGCTCAAAACAAAACCTCGAAAAGCCCTGCGGTCTTTTTT 120
121 TTTTTTTTTTTTTTTTTTGGACAGTCTTGCTCTGTCGCCAGGCTGGAGTACAATGGT 180
121 TTTTTTTTTTTTTTTTTTGGACAGTCTTGCTCTGTCGCCAGGCTGGAGTACAATGGT 180
181 CGATCTTGGCTCACTGCAACCTCTGCTCCCAAGTTCAGAAATCTTCTGCTCAGCC 240
181 CGATCTTGGCTCACTGCAACCTCTGCTCCCAAGTTCAGAAATCTTCTGCTCAGCC 240
241 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTAGTAGAGCGGGTTTTCAC 300
241 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTAGTAGAGCGGGTTTTCAC 300
301 CATGTTGTCAGGCTGTCNGAATCTCTGCTCAGCTGATCCACCGCTCGGCCCC 360
301 CATGTTGTCAGGCTGTCNGAATCTCTGCTCAGCTGATCCACCGCTCGGCCCC 360
361 CAAAGTACTAGGATACAGCGGTGAGCCACCGCTGAGCGCCCTGCGGTCTTTAATCA 420
361 CAAAGTACTAGGATACAGCGGTGAGCCACCGCTGAGCGCCCTGCGGTCTTTAATCA 420
421 AGTAAAGTGCATTAATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
421 AGTAAAGTGCATTAATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
481 AAATGCAATCNCCTTATTAGTGTAGGAAACAGATCTCAAAACAGAGTTTGTGACAG 540
481 AAATGCAATCNCCTTATTAGTGTAGGAAACAGATCTCAAAACAGAGTTTGTGACAG 540
541 ACCGAGGAAACCTGGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
541 ACCGAGGAAACCTGGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
601 TTCCAAAGGCGAGTCTTCCCGGACCGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
601 TTCCAAAGGCGAGTCTTCCCGGACCGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
661 AGACTCTCAGTGTGGCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
661 AGACTCTCAGTGTGGCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
721 CCCTGGCCAGCTCTTAGAGCGGGGCGCGCGCCCGCGAGAGAGAGAGAGAGAGAGAG 780
721 CCCTGGCCAGCTCTTAGAGCGGGGCGCGCGCCCGCGAGAGAGAGAGAGAGAGAGAG 780
781 GAGCGGGGAGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
781 GAGCGGGGAGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
841 CTTGCTCTCGTCCGCGCGCCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900

Db 841 CTTGCTCTCGTCCGCGCGCCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Qy 901 CTTGCTCTCGTCCGCGCGCCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 CTTGCTCTCGTCCGCGCGCCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Qy 961 TTGCCCCCGCGCGCAGAGAGCGGAGCGCGAGCGCGCTCGCGCAGCGGAGCTGCGAG 1020
Db 961 TTGCCCCCGCGCGCAGAGAGCGGAGCGCGAGCGCGCTCGCGCAGCGGAGCTGCGAG 1020
Qy 1021 TACTTGGGCGAGATCCAAACACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1021 TACTTGGGCGAGATCCAAACACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Qy 1081 ACCGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 1081 ACCGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Qy 1141 GCGCCCTCGCGGAGCGGTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1187
Db 1141 GCGCCCTCGCGGAGCGGTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1187
RESULT 5
US-09-658-659-8
; Sequence 8, Application US/09658659
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: GENE SEQUENCE VARIANCE IN GENES RELATED
; TITLE OF INVENTION: TO FOLATE METABOLISM HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015001
; CURRENT APPLICATION NUMBER: US/09/658,659
; CURRENT FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 18597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 701, 13751
; OTHER INFORMATION: n = c or a
; NAME/KEY: misc_feature
; LOCATION: 716, 1293, 2401, 2429, 2618, 3083, 3125, 3635, 4256, 4898,
; LOCATION: 5062, 5167, 11069, 13298, 14479, 14730, 14796, 15344, 15450,
; LOCATION: 15503, 15590, 15840, 16149
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc_feature
; LOCATION: 732, 1379, 1590, 2488, 3212, 5006, 11238, 11422, 11686,
; LOCATION: 12598, 13171, 13645, 13782, 13806, 13813, 14586, 14788,
; LOCATION: 15042, 15546, 15770
; OTHER INFORMATION: n = c or t
; NAME/KEY: misc_feature
; LOCATION: 1322, 1688
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 2594, 11293, 16199, 16203
; OTHER INFORMATION: n = g or t
; NAME/KEY: misc_feature
; LOCATION: 3619
; OTHER INFORMATION: n = a or t
; NAME/KEY: misc_feature

LOCATION: 14547
OTHER INFORMATION: nucleotide in position 14547 is t, or absent
US-09-658-659-8

Query Match 68.3%; Score 811; DB 25; Length 18597;
Best Local Similarity 99.5%; Pred. No. 8.8e-93;
Matches 1181; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 GATCGCGCCACTGACCTCCAGCTGGGTGAGAGAGAGAGACTCTGTCTCAAAAAA 60
DB 78 GATCGCGCCACTGACCTCCAGCTGGGTGAGAGAGAGAGACTCTGTCTCAAAAAA 137
QY 61 AAAAAAGCCGCGAGGCTCAAAACAAAAACCTCGGAAAGCCCTGGGGGCTTTTTTTT 120
DB 138 AAAAAAGCCGCGAGGCTCAAAACAAAAACCTCGGAAAGCCCTGGGGGCTTTTTTTT 197
QY 121 TTTTTTTTTTTTTTTTTTGGGACAGTCTTGCTCTGTCGCCAGGCTGGAGTACAATGT 180
DB 198 TTTTTTTTTTTTTTTTTTGGGACAGTCTTGCTCTGTCGCCAGGCTGGAGTACAATGT 257
QY 181 CGAATCTGGCTCACTGCAACTCTGCTCCAGGTTCAAGCAATTTCTTCTGCTCAGCC 240
DB 258 CGAATCTGGCTCACTGCAACTCTGCTCCAGGTTCAAGCAATTTCTTCTGCTCAGCC 317
QY 241 TCCCAAGTAGCCACCAAGCCAGCTAAATTTTTTGTANTTTTGTAGAGACGGGGTTTAC 300
DB 318 TCCCAAGTAGCCACCAAGCCAGCTAAATTTTTTGTANTTTTGTAGAGACGGGGTTTAC 377
QY 301 CATGTGTCCAGGCTGCTGTAACCTCTGACTCAGGTGATCCACCGGCTCGGCCCCC 360
DB 378 CATGTGTCCAGGCTGCTGTAACCTCTGACTCAGGTGATCCACCGGCTCGGCCCCC 437
QY 361 CAAAGTACTAGGATTACAGGCTGAGCCACCGGCTCAGCGCCCTGCGGGTTTTAATCA 420
DB 438 CAAAGTACTAGGATTACAGGCTGAGCCACCGGCTCAGCGCCCTGCGGGTTTTAATCA 497
QY 421 AGTAGAAAAGCTGCATTAACACCTGCTTCGTTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 498 AGTAGAAAAGCTGCATTAACACCTGCTTCGTTGCTGCTGCTGCTGCTGCTGCTGCT 556
QY 481 AATGCAATATCTTATAGTTAGTAAACAGATCTCAACAGAGCTTTGTGTGCAAG 540
DB 557 AATGCAATATCTTATAGTTAGTAAACAGATCTCAACAGAGCTTTGTGTGCAAG 616
QY 541 ACCGCGAGAAACGTTGGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 617 ACCGCGAGAAACGTTGGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
QY 601 TTCCCAAGGCGGAGTCTCTCCNGCCACCGCACTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 677 TTCCCAAGGCGGAGTCTCTCCNGCCACCGCACTGCTGCTGCTGCTGCTGCTGCTGCT 736
QY 661 AAGACTCTAGCTGTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 737 AAGACTCTAGCTGTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
QY 721 CCCTGGCGACGCTCTTAGAGCGGGGGCGCGGACCCCGCGAGCAGGAAGAGGCG 780
DB 797 CCCTGGCGACGCTCTTAGAGCGGGGGCGCGGACCCCGCGAGCAGGAAGAGGCG 856
QY 781 GAGCGGGGAGCGCGCGGAAAGCGCGGAGGGGTCTGCGACCGCGCCACTTGG 840
DB 857 GAGCGGGGAGCGCGCGGAAAGCGCGGAGGGGTCTGCGACCGCGCCACTTGG 916
QY 841 CTTGCTCCGTCGCGCGGCGGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 917 CTTGCTCCGTCGCGCGGCGGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
QY 901 CTTGCTCCGTCGCGCGGCGGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
DB 977 CTTGCTCCGTCGCGCGGCGGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
QY 961 TTGCCCCCCCCCGCCAGCAGAGCGGGAACGCGAGCCGCTGCTGCGCGGAGCTGCAG 1020

DB 1037 TTGCCCCCCCCCGCCAGCAGAGCGGGAACCGGCTGCGCGGAGCTGCAG 1096
QY 1021 TACTTGGGCGAGATCAACACATCTTCCGCTGCGGCGTCAAGAAAGGACGACCGGCG 1080
DB 1097 TACTTGGGCGAGATCAACACATCTTCCGCTGCGGCGTCAAGAAAGGACGACCGGCG 1156
QY 1081 ACCGCGACCTTGTGGTATTCGGCTATGCGGCTGCGGCTGAGAGCTGAGGCTGAGCGCG 1140
DB 1157 ACCGCGACCTTGTGGTATTCGGCTATGCGGCTGCGGCTGAGAGCTGAGGCTGAGCGCG 1216
QY 1141 GCGCCCTGCGGAGCGGCTGCGGAGGAGGAGGCGGCGGCTGAGGGA 1187
DB 1217 GCGCCCTGCGGAGCGGCTGCGGAGGAGGAGGCGGCGGCTGAGGGA 1263

RESULT 6
US-09-962-665-8
; Sequence 8, Application US/09962665
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: POLY(POLYGLUTAMATE SYNTHETASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015004
; CURRENT APPLICATION NUMBER: US/09/962,665
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 18597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 701, 13751
; OTHER INFORMATION: n = c or a
; NAME/KEY: misc_feature
; LOCATION: 716, 1293, 2401, 2429, 2618, 3083, 3125, 3635, 4256, 4898,
; LOCATION: 5062, 5167, 11069, 13298, 14479, 14730, 14796, 15344, 15450,
; LOCATION: 15503, 15590, 15840, 16149
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc_feature
; LOCATION: 732, 1379, 1590, 2488, 3212, 5006, 11238, 11422, 11686,
; LOCATION: 12598, 13171, 13645, 13782, 13806, 13813, 14586, 14788,
; LOCATION: 15042, 15546, 15770
; OTHER INFORMATION: n = c or t
; NAME/KEY: misc_feature
; LOCATION: 1322, 1688
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 2594, 11293, 16199, 16203
; OTHER INFORMATION: n = g or t
; NAME/KEY: misc_feature
; LOCATION: 3619
; OTHER INFORMATION: n = a or t
; NAME/KEY: misc_feature
; LOCATION: 14547
; OTHER INFORMATION: nucleotide in position 14547 is t, or absent
US-09-962-665-8

Query Match 68.3%; Score 811; DB 36; Length 18597;
Best Local Similarity 99.5%; Pred. No. 8.8e-93;

Matches 1181; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
2Y 1 GATCGCGCACTGCACTCCAGCTCGGTGAGAGCGGAGACTCTGTCTCAAAAAA 60
Db 78 GATCGCGCACTGCACTCCAGCTCGGTGAGAGCGGAGACTCTGTCTCAAAAAA 137
2Y 61 AAAAAGACCGCAGGCTCAACAAAAAAGCTCGGAAAGCTTCGCGGTCTTTT 120
Db 138 AAAAAGACCGCAGGCTCAACAAAAAAGCTCGGAAAGCTTCGCGGTCTTTT 197
2Y 121 TTTTCTTTTCTTTTCTTTTGGACAGCTCTGCTCTGCGCAGGCTGAGTCAATG 180
Db 198 TTTTCTTTTCTTTTCTTTTGGACAGCTCTGCTCTGCGCAGGCTGAGTCAATG 257
2Y 181 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGCTCAAGCAATTTCTTCTCAGCC 240
Db 258 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGCTCAAGCAATTTCTTCTCAGCC 317
2Y 241 TCCAACTAGGACACGCGGCTCAATTTTGTANTTTTGTAGAGACGGGTTTAC 300
Db 318 TCCAACTAGGACACGCGGCTCAATTTTGTACTTTTGTAGAGACGGGTTTAC 377
2Y 301 CATGTGTGTCAGGCTGTGAACTGCTGACCTCAGGTATCCACCCGCTCGCC 360
Db 378 CATGTGTGTCAGGCTGTGAACTGCTGACCTCAGGTATCCACCCGCTCGCC 437
2Y 361 CAAAGTACTAGGATTAAGCGCTGAGCCAGCCGCTCCAGCGCTCGGCGTTTAA 420
Db 438 CAAAGTACTAGGATTAAGCGCTGAGCCAGCCGCTCCAGCGCTCGGCGTTTAA 497
2Y 421 AGTAGAAGCTGATATACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 498 AGTAGAAGCTGATATACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556
2Y 481 AAATCCAAATCCTTATAGTTGTTAGGAAACAGATCTCAACAGCAGTTTGT 540
Db 557 AAATCCAAATCCTTATAGTTGTTAGGAAACAGATCTCAACAGCAGTTTGT 616
2Y 541 ACCGAGGAAACGTTGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 617 ACCGAGGAAACGTTGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676
2Y 601 TTCCAAAGGCGGAGTCTTCCGCGGACGAGCTGCTGCTGCTGCTGCTGCTG 660
Db 677 TTCCAAAGGCGGAGTCTTCCGCGGACGAGCTGCTGCTGCTGCTGCTGCTG 736
2Y 661 AAGACTCTCAGCTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 737 AAGACTCTCAGCTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
2Y 721 CCGCTGCGCACGCTCTCTAGAGCGGCGCGCGGACCCCGCGGAGCAGAGAG 780
Db 797 CCGCTGCGCACGCTCTCTAGAGCGGCGCGCGGACCCCGCGGAGCAGAGAG 856
2Y 781 GAGCGCGGAGCGCGCGGAGAAAGCGCGGAGAGAGAGAGAGAGAGAGAG 840
Db 857 GAGCGCGGAGCGCGCGGAGAAAGCGCGGAGAGAGAGAGAGAGAGAGAG 916
2Y 841 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 917 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
2Y 901 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 977 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1036
2Y 961 TTGCCCCCGCGCGCAGAGCGGAGCGCGCGGCTGCTGCTGCTGCTGCTGCT 1020
Db 1037 TTGCCCCCGCGCGCAGAGCGGAGCGCGCGGCTGCTGCTGCTGCTGCTGCT 1096
2Y 1021 TACCTGGGAGAGTCAACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1097 TACCTGGGAGAGTCAACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1156

QY 1081 ACCGCGACCTCTCGTATTCGGCATGCGGCGCTACAGCTGAGAGTGAAGCGCG 1140
Db 1157 ACCGCGACCTCTCGTATTCGGCATGCGGCGCTACAGCTGAGAGTGAAGCGCG 1216
QY 1141 GCGCCCTCGGAGCGGTGCGGGAAGGAGGCGCGGCTGGGA 1187
Db 1217 GCGCCCTCGGAGCGGTGCGGGAAGGAGGCGCGGCTGGGA 1263

RESULT 7

US-09-962-677-8
; Sequence 8, Application US/09962677
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
; TITLE OF INVENTION: THE TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015003
; CURRENT APPLICATION NUMBER: US/09/962.677
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 18597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 701..13751
; OTHER INFORMATION: n = c or a
; NAME/KEY: misc feature
; LOCATION: 716..1293, 2401..2429, 2618, 3083, 3125, 3635, 4256, 4898,
; LOCATION: 5062, 5167, 11069, 13298, 14479, 14730, 14796, 15344, 15450,
; LOCATION: 15503, 15590, 15940, 16149
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc feature
; LOCATION: 732..1379, 1590, 2488, 3212, 5006, 11238, 11422, 11686,
; LOCATION: 12598, 13171, 13645, 13782, 13806, 13813, 14586, 14788,
; LOCATION: 15042, 15546, 15770
; OTHER INFORMATION: n = c or t
; NAME/KEY: misc feature
; LOCATION: 1322..1688
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc feature
; LOCATION: 2594, 11293, 16199, 16203
; OTHER INFORMATION: n = g or t
; NAME/KEY: misc feature
; LOCATION: 3619
; OTHER INFORMATION: n = a or t
; NAME/KEY: misc feature
; LOCATION: 14547
; OTHER INFORMATION: nucleotide in position 14547 is t, or absent
US-09-962-677-8

Query Match 68.3%; Score 811; DB 36; Length 18597;
Best Local Similarity 99.5%; Pred. No. 8.8e-93;
Matches 1181; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 GATCGCGCACTGCACTCCAGCTCGGTGAGAGCGGAGACTCTGTCTCAAAAAA 60
Db 78 GATCGCGCACTGCACTCCAGCTCGGTGAGAGCGGAGACTCTGTCTCAAAAAA 137

198 TTTTCTTTTCTTTTCTTTTGGGACAGTCTTCTGCTGCGCCAGGCTGAGTACAATGGT 257
181 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGAAATCTTCTGCTCAGCC 240
258 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGAAATCTTCTGCTCAGCC 317
241 TCCCAAGTACGACCAACGCGGCTGAGTAAATTTTGTATTTTGTAGAGAGGGGTTTCA 300
318 TCCCAAGTACGACCAACGCGGCTGAGTAAATTTTGTATTTTGTAGAGAGGGGTTTCA 377
301 CATGTTGTCAGGCTGCTGCTGAGTAAATTTTGTATTTTGTAGAGAGGGGTTTCA 360
378 CATGTTGTCAGGCTGCTGCTGAGTAAATTTTGTATTTTGTAGAGAGGGGTTTCA 437
361 CAAAGTACTAGGATTAACAGGCTGAGCAACGCGCTGAGGATCCAGCCGCTCGGCCCC 420
438 CAAAGTACTAGGATTAACAGGCTGAGCAACGCGCTGAGGATCCAGCCGCTCGGCCCC 497
421 AGTAGAAGAGCTCAATTAACAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
498 AGTAGAAGAGCTCAATTAACAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556
481 AAATGCAAAATCCTTTATTTAGTTTGTAGAAAAGAGTCAAAAGAGGTTTGTGACAAG 540
557 AAATGCAAAATCCTTTATTTAGTTTGTAGAAAAGAGTCAAAAGAGGTTTGTGACAAG 616
541 ACCGACGAGAAAAGTGGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
617 ACCGACGAGAAAAGTGGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
601 TTCCCAAGAGGCGGAGTCTTCCGCGGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
677 TTCCCAAGAGGCGGAGTCTTCCGCGGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
661 AAGACTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
737 AAGACTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
721 CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGAGCCCGCGAGCAGGAGAGGCG 780
797 CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGAGCCCGCGAGCAGGAGAGGCG 856
781 GAGCGCGGAGCGCGCGGAGAAAGGCGCGGAGAGGAGGCTCTGCTGCTGCTGCTGCTGCT 840
857 GAGCGCGGAGCGCGCGGAGAAAGGCGCGGAGAGGAGGCTCTGCTGCTGCTGCTGCTGCT 916
841 CTTGCTCTCGTCCGCGCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
917 CTTGCTCTCGTCCGCGCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
901 CTTGCTCTCGTCCGCGCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
977 CTTGCTCTCGTCCGCGCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
961 TTGCGCGCGGCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
1037 TTGCGCGCGGCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1096
1021 TACCTGGGCGAGTCCAAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
1097 TACCTGGGCGAGTCCAAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156
1081 ACCGCGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
1157 ACCGCGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1216
1141 GGCCTCTGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1187
1217 GGCCTCTGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1263

RESULT 9

US-09-577-266-11

; Sequence 11, Application US/09577266
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Stenroos, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; TITLE OF INVENTION: DEVELOPMENTAL DISORDERS
; FILE REFERENCE: 601-1-057N
; CURRENT APPLICATION NUMBER: US/09/577,266
; CURRENT FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/136,198
; PRIOR FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 11
; LENGTH: 18596
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-577-266-11

Query Match 55.4%; Score 658; DB 22; Length 18596;
Best Local Similarity 99.2%; Pred. No. 7.7e-74;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1 GATCGCGCACTGCACTCCAGCGCTGGGTGAGAGAGCGAGTCTGCTCAAAAAA 60
DB 78 GATCGCGCACTGCACTCCAGCGCTGGGTGAGAGAGCGAGTCTGCTCAAAAAA 137
QY 61 AAAAAAGCGCGAGGCTCAAAAAAAGCTCGGAAAGCCCTGGCGGCTCTTTT 120
DB 138 AAAAAAGCGCGAGGCTCAAAAAAAGCTCGGAAAGCCCTGGCGGCTCTTTT 197
QY 121 TTTTCTTTTCTTTTCTTTTGGGACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
DB 198 TTTTCTTTTCTTTTCTTTTGGGACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 257
QY 181 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 240
DB 258 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 317
QY 241 TCCCAAGTACGACCAACGCGGCTGAGTAAATTTTGTATTTTGTAGAGAGGGGTTTCA 300
DB 318 TCCCAAGTACGACCAACGCGGCTGAGTAAATTTTGTATTTTGTAGAGAGGGGTTTCA 377
QY 301 CATGTTGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 378 CATGTTGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 437
QY 361 CAAAGTACTAGGATTAACAGGCTGAGCAACGCGCTGAGAGAGGCGGCTGAGAGAGGCG 420
DB 438 CAAAGTACTAGGATTAACAGGCTGAGCAACGCGCTGAGAGAGGCGGCTGAGAGAGGCG 497
QY 421 AGTAGAAGAGCTCAATTAACAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 498 AGTAGAAGAGCTCAATTAACAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556
QY 481 AAATGCAAAATCCTTTATTTAGTTTGTAGAAAAGAGTCAAAAGAGGTTTGTGACAAG 540
DB 557 AAATGCAAAATCCTTTATTTAGTTTGTAGAAAAGAGTCAAAAGAGGTTTGTGACAAG 616
QY 541 ACCGACGAGAAAAGTGGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 617 ACCGACGAGAAAAGTGGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
QY 601 TTCCCAAGAGGCGGAGTCTTCCGCGGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 677 TTCCCAAGAGGCGGAGTCTTCCGCGGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
QY 661 AAGACTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 737 AAGACTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
QY 721 CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGAGCCCGCGAGCAGGAGAGGCG 780

RESULT 11
US-09-880-107-1590
Sequence 1590, Application US/09880107
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1590
LENGTH: 18596
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. D00596
US-09-880-107-1590
Query Match 55.4%; Score 658; DB 33; Length 18596;
Best Local Similarity 99.2%; Pred. No. 7.7e-74;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 1 GATCGCGCCACTGCACTCCAGCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAA 60
DB 78 GATCGCGCCACTGCACTCCAGCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAA 137
QY 61 AAAAGACCGCGCGGCTCAACAAAAAAGCTCGGAAAGCCCTGCGGTCTTTT 120
DB 138 AAAAGACCGCGCGGCTCAACAAAAAAGCTCGGAAAGCCCTGCGGTCTTTT 197
QY 121 TTTTCTTTTCTTTTCTTTTGGGACAGCTCTGTCTGCGCCAGGCTGGAGTACATG 180
DB 198 TTTTCTTTTCTTTTCTTTTGGGACAGCTCTGTCTGCGCCAGGCTGGAGTACATG 257
QY 181 CGGACTTGCTCACTCAACCTCTGCTCCAGCTTCAAGCAATTTCTGCTCAGCC 240
DB 258 CGGACTTGCTCACTCAACCTCTGCTCCAGCTTCAAGCAATTTCTGCTCAGCC 317
QY 241 TCCCAAGTAGCCACACCGCCAGCTAATTTTGTANTTTTGTAGTACAGCGGGTTT 300
DB 318 TCCCAAGTAGCCACACCGCCAGCTAATTTTGTANTTTTGTAGTACAGCGGGTTT 377
QY 301 CATGTTGTCAGGCTGTCTGTAACCTCTGACTCAGTGATCAACCGCTCGGCCCC 360
DB 378 CATGTTGTCAGGCTGTCTGTAACCTCTGACTCAGTGATCAACCGCTCGGCCCC 437
QY 361 CAAAGTACTAGGATACAGGCGTACAGCCCGGTCAGCGCCCTGCGGTTTAAATCA 420
DB 438 CAAAGTACTAGGATACAGGCGTACAGCCCGGTCAGCGCCCTGCGGTTTAAATCA 497
QY 421 AGTAGAAAAGCTGATTAACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 498 AGTAGAAAAGCTGATTAACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 556
QY 481 AATGCAATCTTATTTAGTTAGTAAACAGATCTCAACAGCAGTTTGTGNGACAAG 540
DB 557 AATGCAATCTTATTTAGTTAGTAAACAGATCTCAACAGCAGTTTGTGNGACAAG 616
QY 541 ACCCGAGAAACCTGGGAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 617 ACCCGAGAAACCTGGGAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676
QY 601 TTTCCAAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 677 TTTCCAAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736

QY 661 AAGACTCTCAGTGTGGCCCTGGCTCCGTTCTGTGTCACACCCGCTGCTCTGCGTTTC 720
DB 737 AAGACTCTCAGTGTGGCCCTGGCTCCGTTCTGTGTCACACCCGCTGCTCTGCGTTTC 796
QY 721 CCCCTGGCCACGCTCTCTAGAGCGGGGCGCGCGAGACCCCGCGAGAGAGGCG 780
DB 797 CCCCTGGCCACGCTCTCTAGAGCGGGGCGCGCGAGACCCCGCGAGAGAGGCG 856
QY 781 GAGCGGGGACGCGCGCGGAAAAAGCGCGGAGAGGGTCTTGCACCGCGCCACTTGG 840
DB 857 GAGCGGGGACGCGCGCGGAAAAAGCGCGGAGAGGGTCTTGCACCGCGCCACTTGG 916
QY 841 CCTGCTCCGTTCCGCGCGCGCACTTGGCTTGCCTCCGTCGCGCGCGCACTTTCGCTG 900
DB 917 CCTGCTCCGTTCCGCGCGCGCACTTGGCTTGCCTCCGTCGCGCGCGCACTTTCGCTG 976
QY 901 CCTGCTCCGTTCCGCGCGCGCACTTGGCTTGCCTCCGTCGCGCGCGCACTTTCGCTG 960
DB 977 CCTGCTCCGTTCCGCGCGCGCACTTGGCTTGCCTCCGTCGCGCGCGCACTTTCGCTG 1036
QY 961 TTGCCCCCGCCGCGAGCGGAGCGCGCGCGCTCCGCGCGAGCGGGAGCTGCAG 1020
DB 1037 TTGCCCCCGCCGCGAGCGGAGCGCGCGCGCTCCGCGCGAGCGGGAGCTGCAG 1096
QY 1021 TACCTGGGAGATCCAAACACATCTCTGCTGCGCGCTCAGGAAGACGACCGCACGGGC 1080
DB 1097 TACCTGGGAGATCCAAACACATCTCTGCTGCGCGCTCAGGAAGACGACCGCACGGGC 1156
QY 1081 ACCGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
DB 1157 ACCGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1216
QY 1141 GCGCCCTCGCGGAGCGGTCGCGGAGAGGAGGAGCGCGCTGGGA 1187
DB 1217 GCGCCCTCGCGGAGCGGTCGCGGAGAGGAGGAGCGCGCTGGGA 1263

RESULT 12
US-09-954-531-124
Sequence 124, Application US/09954531
GENERAL INFORMATION:
APPLICANT: Weaver, Zoe
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
TITLE OF INVENTION: Gene Sets
FILE REFERENCE: 689290-77
CURRENT APPLICATION NUMBER: US/09/954,531
CURRENT FILING DATE: 2002-05-02
PRIOR APPLICATION NUMBER: US/60/233,133
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,009
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,034
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,509
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: US/60/234,567
PRIOR FILING DATE: 2000-09-22
NUMBER OF SEQ ID NOS: 1392
SOFTWARE: PatentIn version 3.0
SEQ ID NO 124
LENGTH: 18596
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-531-124
Query Match 55.4%; Score 658; DB 36; Length 18596;
Best Local Similarity 99.2%; Pred. No. 7.7e-74;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 1 GATCGCGCCACTGCACTCCAGCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAA 60
DB 78 GATCGCGCCACTGCACTCCAGCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAA 137

QY 61 AAAAAAGCGCGGCTCAACAAAAACCTCGAAAAAGCCCTGGGGGCTTTT 120
Db 138 AAAAAAGCGCGGCTCAACAAAAACCTCGAAAAAGCCCTGGGGGCTTTT 197
QY 121 TTTTATTTTATTTTATTTTGGGACAGTCTTGCTCTGTCCGCCAGCTGGAGTACAATGGT 180
Db 198 TTTTATTTTATTTTATTTTGGGACAGTCTTGCTCTGTCCGCCAGCTGGAGTACAATGGT 257
QY 181 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTTCTGCTCAGCC 240
Db 258 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTTCTGCTCAGCC 317
QY 241 TCCCAAGTAGGCTACAGCCAGCTTAATTTTGTATTTAGTAGAGAGCGGGGTTTAC 300
Db 318 TCCCAAGTAGGCTACAGCCAGCTTAATTTTGTATTTAGTAGAGAGCGGGGTTTAC 377
QY 301 CATGTTGTCCAGGCTGTTGAACTCTGACCTCAGGTGATCCACCCGCTCGGCCCCC 360
Db 378 CATGTTGTCCAGGCTGTTGAACTCTGACCTCAGGTGATCCACCCGCTCGGCCCCC 437
QY 361 CAAGTACTAGATTTACAGGCTGAGCCAGCTGCTCAGTCTGAGTCAAGCAATGG 480
Db 438 CAAGTACTAGATTTACAGGCTGAGCCAGCTGCTCAGTCTGAGTCAAGCAATGG 540
QY 421 AGTAGAAAAGCTGCAATTAATTTTGTATTTAGTAGAGAGCGGGGTTTAC 480
Db 498 AGTAGAAAAGCTGCAATTAATTTTGTATTTAGTAGAGAGCGGGGTTTAC 556
QY 481 AAATGCAAAATCNCCTTATTTAGTTAGTGAAGCAAGATCTCAAAAGCAAGTGTGNGACAAG 540
Db 557 AAATGCAAAATCNCCTTATTTAGTTAGTGAAGCAAGATCTCAAAAGCAAGTGTGNGACAAG 616
QY 541 ACCGAGGAAAAGCTGGGAACCTGCTGCTGCTGAGAGAGCGGCTCGACCAAGCG 600
Db 617 ACCGAGGAAAAGCTGGGAACCTGCTGCTGCTGAGAGAGCGGCTCGACCAAGCG 676
QY 601 TTCCCAAGGCGGAGTCTTCCGNGCCACCGGACCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 677 TTCCCAAGGCGGAGTCTTCCGNGCCACCGGACCTGCTGCTGCTGCTGCTGCTGCTGCT 736
QY 661 AAGACTCAGCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 737 AAGACTCAGCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
QY 721 CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGGACCCCGGACCGGCGGAGAGCG 780
Db 797 CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGGACCCCGGACCGGCGGAGAGCG 856
QY 781 GAGCGGGAAGCGCGGGAAGGCGGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Db 857 GAGCGGGAAGCGCGGGAAGGCGGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 916
QY 841 CTTGCTCTCCGTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 917 CTTGCTCTCCGTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
QY 901 CTTGCTCTCCGTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 977 CTTGCTCTCCGTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
QY 961 TTGCCCCCGCGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
Db 1037 TTGCCCCCGCGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1096
QY 1021 TACCTGGGAGATCCAAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1097 TACCTGGGAGATCCAAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156
QY 1081 ACCGGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db 1157 ACCGGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1216

QY 1141 GGCCCTCGCGACCGGTTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1187
Db 1217 GGCCCTCGCGACCGGTTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1263

RESULT 13
US-09-954-531-348
; Sequence 348, Application US/09954531
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; TITLE OF INVENTION: Gene Sets
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 348
; LENGTH: 18596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-348

Query Match 55.4%; Score 658; DB 36; Length 18596;
Best Local Similarity 99.2%; Pred. No. 7,7e-74;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1 GATCGCGCACTGCACTCCAGGCTGGGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
Db 78 GATCGCGCACTGCACTCCAGGCTGGGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 137
QY 61 AAAAAAGCGCGGCTCAACAAAAACCTCGAAAAAGCCCTGGGGGCTTTT 120
Db 138 AAAAAAGCGCGGCTCAACAAAAACCTCGAAAAAGCCCTGGGGGCTTTT 197
QY 121 TTTTATTTTATTTTATTTTGGGACAGTCTTGCTCTGTCCGCCAGCTGGAGTACAATGGT 180
Db 198 TTTTATTTTATTTTATTTTGGGACAGTCTTGCTCTGTCCGCCAGCTGGAGTACAATGGT 257
QY 181 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTTCTGCTCAGCC 240
Db 258 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTTCTGCTCAGCC 317
QY 241 TCCCAAGTAGGCTACAGCCAGCTTAATTTTGTATTTAGTAGAGAGCGGGGTTTAC 300
Db 318 TCCCAAGTAGGCTACAGCCAGCTTAATTTTGTATTTAGTAGAGAGCGGGGTTTAC 377
QY 301 CATGTTGTCCAGGCTGTTGAACTCTGACCTCAGGTGATCCACCCGCTCGGCCCCC 360
Db 378 CATGTTGTCCAGGCTGTTGAACTCTGACCTCAGGTGATCCACCCGCTCGGCCCCC 437
QY 361 CAAGTACTAGATTTACAGGCTGAGCCAGCTGCTCAGTCTGAGTCAAGCAATGG 480
Db 438 CAAGTACTAGATTTACAGGCTGAGCCAGCTGCTCAGTCTGAGTCAAGCAATGG 540
QY 421 AGTAGAAAAGCTGCAATTAATTTTGTATTTAGTAGAGAGCGGGGTTTAC 480
Db 498 AGTAGAAAAGCTGCAATTAATTTTGTATTTAGTAGAGAGCGGGGTTTAC 556
QY 481 AAATGCAAAATCNCCTTATTTAGTTAGTGAAGCAAGATCTCAAAAGCAAGTGTGNGACAAG 540
Db 557 AAATGCAAAATCNCCTTATTTAGTTAGTGAAGCAAGATCTCAAAAGCAAGTGTGNGACAAG 616

QY	1	GATCGCCCACTGCACCTCAGCTGGGTGAGAGCGAGACTCTCTCTCAAAAAA	60
Db	78	GATCGCCCACTGCACCTCAGCTGGGTGAGAGCGAGACTCTCTCTCAAAAAA	137
QY	61	AAAAAGACCGCCAGGCTCAAAACAAAAA	120
Db	138	AAAAAGACCGCCAGGCTCAAAACAAAAA	197
QY	121	TT	180
Db	198	TT	257
QY	181	CGGATCTTGGCTCACTGCAACCTCCAGCTTCAAGCAATCTTCTGCTCAGCC	240
Db	258	CGGATCTTGGCTCACTGCAACCTCCAGCTTCAAGCAATCTTCTGCTCAGCC	317
QY	241	TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTTGTAGAGCGGGTTTCA	300
Db	318	TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTTGTAGAGCGGGTTTCA	377
QY	301	CATGTTGTCAGGCTGCTCTGAACTCTCTGACCTCAGGTATCAACCGCTCGGCC	360
Db	378	CATGTTGTCAGGCTGCTCTGAACTCTCTGACCTCAGGTATCAACCGCTCGGCC	437
QY	361	CAAGTACTTAGGATTTACAGCGTGAAGCAACCGCTGAGCGCTTGTAAATCA	420
Db	438	CAAGTACTTAGGATTTACAGCGTGAAGCAACCGCTGAGCGCTTGTAAATCA	497
QY	421	AGTAGAAAGCTGATTAACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTG	480
Db	498	AGTAGAAAGCTGATTAACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTG	556
QY	481	AAATGCAATCCTTTTATTTAGTTAGTAAACAGATCTCAAAACAGCAGTTT	540
Db	557	AAATGCAATCCTTTTATTTAGTTAGTAAACAGATCTCAAAACAGCAGTTT	616
QY	541	ACCGGAGAAACGTTGGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	600
Db	617	ACCGGAGAAACGTTGGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	676
QY	601	TTCCCAAGGCGCAGTCTCTCCCGCCACCGCACTGCTGCTGCTGCTGCTGCT	660
Db	677	TTCCCAAGGCGCAGTCTCTCCCGCCACCGCACTGCTGCTGCTGCTGCTGCT	736
QY	661	AAGACTCTCAGCTGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT	720
Db	737	AAGACTCTCAGCTGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT	796
QY	721	CCCTGCGCAGCCTCTTAGAGCGGGGCGCGCGGAGCGCGCGGAGAGAGGCG	780
Db	797	CCCTGCGCAGCCTCTCTTAGAGCGGGGCGCGCGGAGCGCGCGGAGAGAGGCG	856
QY	781	GAGCGGCGGAGCGCGCGGAGAGAGCGCGCGGAGAGAGCGCGCGGAGAGAGG	840
Db	857	GAGCGGCGGAGCGCGCGGAGAGAGCGCGCGGAGAGAGCGCGCGGAGAGAGG	916
QY	841	CCCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	900
Db	917	CCCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	976
QY	901	CCCTGCTCCG	960
Db	977	CCCTGCTCCG	1036
QY	961	TTGCCCCCG	1020
Db	1037	TTGCCCCCG	1096
QY	1021	TACCTGGGCGAGATCCAAACATCTCTCGCTGCTGCTGCTGCTGCTGCTG	1080
Db	1097	TACCTGGGCGAGATCCAAACATCTCTCGCTGCTGCTGCTGCTGCTGCTG	1156
QY	1081	ACCGGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1140

QY	541	ACCAGAGAAAACGTTGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	600
Db	617	ACCAGAGAAAACGTTGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	676
QY	601	TTTCCAAAGGCGCAGTCTTCCCGCCAGCAGCTGCTGCTGCTGCTGCTGCTG	660
Db	677	TTTCCAAAGGCGCAGTCTTCCCGCCAGCAGCTGCTGCTGCTGCTGCTGCTG	736
QY	661	AAGACTCTCAGCTGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT	720
Db	737	AAGACTCTCAGCTGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT	796
QY	721	CCCTGCGCAGCCTCTCTAGAGCGGGGCGCGCGAGACCCCGCGAGAGAGCG	780
Db	797	CCCTGCGCAGCCTCTCTAGAGCGGGGCGCGCGAGACCCCGCGAGAGAGCG	856
QY	781	GAGCGCGGAGCGCGCGGAGAGAGCGCGGAGAGGCTTCTGCAACCGCGCACT	840
Db	857	GAGCGCGGAGCGCGCGGAGAGAGCGCGGAGAGGCTTCTGCAACCGCGCACT	916
QY	841	CCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	900
Db	917	CCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	976
QY	901	CCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	960
Db	977	CCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1036
QY	961	TTGCCCCCG	1020
Db	1037	TTGCCCCCG	1096
QY	1021	TACCTGGGCGAGATCCAAACATCTCTCGCTGCTGCTGCTGCTGCTGCTG	1080
Db	1097	TACCTGGGCGAGATCCAAACATCTCTCGCTGCTGCTGCTGCTGCTGCTG	1156
QY	1081	ACCGGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1140
Db	1157	ACCGGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1216
QY	1141	GGCCCCCGGAGAGGTTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1187
Db	1217	GGCCCCCGGAGAGGTTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1263

RESULT 14
US-09-967-768A-119
Sequence 119, Application US/09967768A
GENERAL INFORMATION:
APPLICANT: Augustus, Meena
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
FILE REFERENCE: 689290-72
CURRENT APPLICATION NUMBER: US/09/967,768A
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US/60/236,109
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,034
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,111
PRIOR FILING DATE: 2000-09-28
NUMBER OF SEQ ID NOS: 325
SOFTWARE: PatentIn version 3.0
SEQ ID NO 119
LENGTH: 18596
TYPE: DNA
ORGANISM: Homo sapiens
US-09-967-768A-119
Query Match 55.4%; Score 658; DB 36; Length 18596;
Best Local Similarity 99.2%; Pred. No. 7.7e-74;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

RESULT 15

Query Match 25.0%; Score 297; DB 1; Length 1536;

Query Match 25.0%; Score 297; DB 1; Length 1536;
Best Local Similarity 100.0%; Pred. No. 9.2e-29;
Matches 297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1073 GCACGGGACCGGACCCCTGCGTATTCGGCATCGAGCGCGCTACAGCTGAGAG 1129

Db 254 GCACGGGACCGGACCCCTGCGTATTCGGCATCGAGCGCGCTACAGCTGAGAG 310

Search completed: November 29, 2002, 05:05:02
Job time : 2485.99 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

```

1 nucleic - nucleic search, using sw model
run on:      November 29, 2002, 02:34:05 ; Search time 413.685 Seconds
              (without alignments)
              1105.026 Million cell updates/sec

```

```

title: US-09-963-333-7
perfect score: 1187
sequence: 1 gatcgcgccactgactcca.....gagggagcgcgactgggga 1187

```

scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

searched: 341543 seqs, 192557720 residues

```
ord size : 15
total number of hits satisfying chosen parameters: 49150
```

```

inimum DB seq length: 0
aximum DB seq length: 200000000

```

Post-processing: Listing first 45 summaries

```

Database : Published Applications NA:*
1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	658	55.4	18596	9	US-09-954-531-124		Sequence 124, App Sequence 348, App
2	658	55.4	18596	9	US-09-954-531-348		Sequence 1590, App
3	658	55.4	18596	10	US-09-880-107-1590		Sequence 1119, App
4	658	55.4	18596	10	US-09-967-768A-119		Sequence 2148, App
5	108	9.1	532	9	US-10-046-935-2148		Sequence 2113, App
C 6	58	4.9	10867	10	US-09-764-869-2113		Sequence 2110, App
C 7	58	4.9	10867	10	US-09-764-877-3970		Sequence 3970, App
C 8	58	4.9	10907	10	US-09-764-869-2112		Sequence 2112, App
C 9	58	4.9	10907	10	US-09-764-877-3968		Sequence 3968, App
C 10	58	4.9	12718	10	US-09-764-877-3972		Sequence 3972, App
C 11	56	4.7	143306	10	US-09-729-920-3		Sequence 3, App
C 12	55	4.6	10894	10	US-09-764-869-2111		Sequence 2111, App
C 13	54	4.5	58837	10	US-09-982-091A-5		Sequence 5, App
C 14	48	4.0	32170	10	US-09-764-860-1108		Sequence 1108, App
C 15	46	3.9	13500	9	US-09-954-531-145		Sequence 145, App
C 16	46	3.9	13500	9	US-09-954-531-363		Sequence 363, App
C 17	46	3.9	13500	10	US-09-962-436-269		Sequence 269, App
C 18	46	3.9	13500	10	US-09-954-456-55		Sequence 55, App
C 19	46	3.9	13500	10	US-09-954-456-1132		Sequence 1132, App

Query Match	55.4%;	Score 636;	DB 5;	Longer
Best Local Similarity	99.2%;	Pred. No. 4.2e-168;		
			0. Mismatches	8;
			Indels	1;
			Caps	1;

1	QY	GATCGGCGCACTGCATCCAGCCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA	60
78	Db	GATCGCGCACTGCATCCAGCCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA	137
61	QY	AAAAAGACCGCCAGGCTCAAAACAAAACCTCGEAAAGCCCTCGCGCTTTTTTT	120
138	Db	AAAAAGACCGCCAGGCTCAAAACAAAACCTCGEAAAGCCCTCGCGCTTTTTTT	197
121	QY	TTTTTTTTTTTTTTTTTTTTTGGGACAGTCTTGCTGTGCGCCAGCGCTGGAGTCAATG	180

ALIGNMENTS

RESULT 1

US-09-954-531-124
; Sequence 124, Application US/09954531

```

; Patent No. US20020165380A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Weaver, Zoe
;
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cancer
;
; TITLE OF INVENTION: Gene Sets
;
; FILE REFERENCE: 689230-77
;
; CURRENT APPLICATION NUMBER: US/09/954,531

```

```

1 FILE REFERENCE: 062230-77
2 CURRENT APPLICATION NUMBER: US/09/954,531
3 CURRENT FILING DATE: 2002-05-02
4 PRIOR APPLICATION NUMBER: US/60/233,133
5 PRIOR FILING DATE: 2000-09-18
6 PRIOR APPLICATION NUMBER: US/60/234,009
7 PRIOR FILING DATE: 2000-09-20
8 PRIOR APPLICATION NUMBER: US/60/234,034
9 PRIOR FILING DATE: 2000-09-20
10 PRIOR APPLICATION NUMBER: US/60/234,509
11 PRIOR FILING DATE: 2000-09-22
12 PRIOR APPLICATION NUMBER: US/60/234,567
13 PRIOR FILING DATE: 2000-09-22
14 NUMBER OF SEQ ID NOS: 1392
15 SOFTWARE: Patent In version 3.0

```

```

; SEQ ID NO 124
; LENGTH: 18596
; TYPE: DNA
; ORGANISM: Homo sapiens
HIS-09-954-531-124

```

Query Match	55.4%;	Score 658;	DB 9;	Length 18596;
Best Local Similarity	99.2%;	Pred. NO. 4.2e-168;		
			Indels 1;	Caps 1;
			8;	

1	QY	GATCGGCGCACTGCATCCAGCCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA	60
78	Db	GATCGCGCACTGCATCCAGCCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA	137
61	QY	AAAAAGACCGCCAGGCTCAAAACA AAAA CTTGGAAGCCCTCGCGCTCTTTTTT	120
138	Db	AAAAAGACCGCCAGGCTCAAAACA AAAA CTTGGAAGCCCTCGCGCTCTTTTTT	197
121	QY	TTTTTTTTTTTTTTTTTTTTTGGACAGTCTTGCTGTGCGCCAGCGCTGGAGTCAATGT	180


```

US-09-954-531-348
; Sequence 348, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C.
; TITLE OF INVENTION: Gene Sets
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 348
; LENGTH: 18596
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-954-531-348

Query Match      55.4%; Score 658; DB 9; Length 18596;
Best Local Similarity 99.2%; Pred. No. 4.2e-168;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1 GATCGCCCACTGCATCCAGCGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 60
Db 78 GATCGCCCACTGCATCCAGCGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 137
QY 61 AAAAGACCGCGAGCGCTCAAAACAAAAACCTCGGAAAAACCTCGGCGGTCTTTT 120
Db 138 AAAAGACCGCGAGCGCTCAAAACAAAAACCTCGGAAAAACCTCGGCGGTCTTTT 197
QY 121 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 180
Db 198 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 257
QY 181 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTTCTCAGCC 240
Db 258 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTTCTCAGCC 317
QY 241 TCCCAAGTAGCACCAACGCGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGTTTCA 300
Db 318 TCCCAAGTAGCACCAACGCGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGTTTCA 377
QY 301 CATGTGTTCAGGGTGGTCTGAACTCTGCTGACCTCAGTGATCCACCGCTCGGCCCCC 360
Db 378 CATGTGTTCAGGGTGGTCTGAACTCTGCTGACCTCAGTGATCCACCGCTCGGCCCCC 437
QY 361 CAAAGTACTAGGATTAACGCGGTGAGCCACCGCTCCAGCGCCCTCGCGGTTTAAATCA 420
Db 438 CAAAGTACTAGGATTAACGCGGTGAGCCACCGCTCCAGCGCCCTCGCGGTTTAAATCA 497
QY 421 AGTAAAAAGCTGCATTATACCACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 498 AGTAAAAAGCTGCATTATACCACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 556
QY 481 AAATCAAAATCNCCTTTTATTTAGTTCTAGGAAACAGATCTCAACAGCAGGTTTCTGACAG 540
Db 557 AAATCAAAATCNCCTTTTATTTAGTTCTAGGAAACAGATCTCAACAGCAGGTTTCTGACAG 616
QY 541 ACCGAGGAAACCTGAGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 617 ACCGAGGAAACCTGAGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676
QY 601 TTCCCAAAAGGCGAGCTCTTCCCGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660

US-09-954-531-348
; Sequence 348, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C.
; TITLE OF INVENTION: Gene Sets
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 348
; LENGTH: 18596
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-954-531-348

Query Match      55.4%; Score 658; DB 9; Length 18596;
Best Local Similarity 99.2%; Pred. No. 4.2e-168;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1 GATCGCCCACTGCATCCAGCGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 60
Db 78 GATCGCCCACTGCATCCAGCGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 137
QY 61 AAAAGACCGCGAGCGCTCAAAACAAAAACCTCGGAAAAACCTCGGCGGTCTTTT 120
Db 138 AAAAGACCGCGAGCGCTCAAAACAAAAACCTCGGAAAAACCTCGGCGGTCTTTT 197
QY 121 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 180
Db 198 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 257
QY 181 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTTCTCAGCC 240
Db 258 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTTCTCAGCC 317
QY 241 TCCCAAGTAGCACCAACGCGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGTTTCA 300
Db 318 TCCCAAGTAGCACCAACGCGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGTTTCA 377
QY 301 CATGTGTTCAGGGTGGTCTGAACTCTGCTGACCTCAGTGATCCACCGCTCGGCCCCC 360
Db 378 CATGTGTTCAGGGTGGTCTGAACTCTGCTGACCTCAGTGATCCACCGCTCGGCCCCC 437
QY 361 CAAAGTACTAGGATTAACGCGGTGAGCCACCGCTCCAGCGCCCTCGCGGTTTAAATCA 420
Db 438 CAAAGTACTAGGATTAACGCGGTGAGCCACCGCTCCAGCGCCCTCGCGGTTTAAATCA 497
QY 421 AGTAAAAAGCTGCATTATACCACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 498 AGTAAAAAGCTGCATTATACCACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 556
QY 481 AAATCAAAATCNCCTTTTATTTAGTTCTAGGAAACAGATCTCAACAGCAGGTTTCTGACAG 540
Db 557 AAATCAAAATCNCCTTTTATTTAGTTCTAGGAAACAGATCTCAACAGCAGGTTTCTGACAG 616
QY 541 ACCGAGGAAACCTGAGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 617 ACCGAGGAAACCTGAGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676
QY 601 TTCCCAAAAGGCGAGCTCTTCCCGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660

```

78 GATCGGCCACTGCACTCCAGCTTCAGGCTGGTGAGAGAGCGAGACTGTCTCAAAAAA 137
61 AAAAAAGCCGCGAGGCTCAAAAAGAAACCTCGGAAAGCCCTGGCGGTCTTTTTTTT 120
138 AAAAAAGCCGCGAGGCTCAAAAAGAAACCTCGGAAAGCCCTGGCGGTCTTTTTTTT 197
121 TT 180
198 TT 257
191 CGGATCTTGGCTCACTGCAAACTCTGCTCCAGGTTCAAGCAATTTCTGCTCAGCC 240
258 CGGATCTTGGCTCACTGCAAACTCTGCTCCAGGTTCAAGCAATTTCTGCTCAGCC 317
241 TCCCAAGTAGCCACCAAGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGGTTTCA 300
318 TCCCAAGTAGCCACCAAGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGGTTTCA 377
301 CATGTTCTCAGGCTCGGTCTGAACTCTGCTCAGGTTCAAGCAATTTCTGCTCAGCC 360
378 CATGTTCTCAGGCTCGGTCTGAACTCTGCTCAGGTTCAAGCAATTTCTGCTCAGCC 437
361 CAAAAGTACTAGGATTAAGGCTGAGCCAGCCGCTCCAGCGCTCGGCGGTTTAAATCA 420
438 CAAAAGTACTAGGATTAAGGCTGAGCCAGCCGCTCCAGCGCTCGGCGGTTTAAATCA 497
421 AGTAGAAAAGCTGATTAATACCACTGCTTCNGTTGCTTCACTGAGAAACGAAATGG 480
498 AGTAGAAAAGCTGATTAATACCACTGCTTCGTTGCTTCACTGAGAAACGAAATGG 556
481 AAATCAATCNCCTTATTTAGTTGTAGAAAACAGATCTCAACAGAGTTTGTGACAG 540
557 AAATCAATCNCCTTATTTAGTTGTAGAAAACAGATCTCAACAGAGTTTGTGACAG 616
541 ACAGCAGAAAACGTTGGGAACTGTCTGCTTGTAGAAAAGCGCGGTGACAGAGG 600
617 ACAGCAGAAAACGTTGGGAACTGTCTGCTTGTAGAAAAGCGCGGTGACAGAGG 676
601 TTCCCAAAAAGGCGCAGTCTTCCGCGCAACCGCACTGCTTCAAGTTCCCGGTTTCT 660
677 TTCCCAAAAAGGCGCAGTCTTCCGCGCAACCGCACTGCTTCAAGTTCCCGGTTTCT 736
661 AAGACTCTCAGTGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTTC 720
737 AAGACTCTCAGTGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTTC 796
721 CCCTCGCGCAAGCTCTCTAGAGCGGGGCGCGCGACCCCGCGCAGAGAGAGGCG 780
797 CCCTCGCGCAAGCTCTCTAGAGCGGGGCGCGCGACCCCGCGCAGAGAGAGGCG 856
781 GAGCGGGAGCGCGCGCGGAAAAGGCGCGCGAGGGGTCTCTGCAACCGCGCACTTGG 840
857 GAGCGGGAGCGCGCGCGGAAAAGGCGCGCGAGGGGTCTCTGCAACCGCGCACTTGG 916
841 CTTGCTCTCGTCCCGCGCGCACTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 900
917 CTTGCTCTCGTCCCGCGCGCACTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 976
901 CTTGCTCTCGTCCCGCGCGCACTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 960
977 CTTGCTCTCGTCCCGCGCGCACTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1036
961 TTGCGCCCCCGCGCGCAGAGAGCGGGAAGCGCGCGGTCCGCGCGAGGAGCTGCAG 1020
1037 TTGCGCCCCCGCGCGCAGAGAGCGGGAAGCGCGCGGTCCGCGCGAGGAGCTGCAG 1096
1021 TACTTGGGCGAGATCCCAACATCTCTGCTGGCGGTCTGAGAGAGCAACCGCACGGGC 1080
1097 TACTTGGGCGAGATCCCAACATCTCTGCTGGCGGTCTGAGAGAGCAACCGCACGGGC 1156
1081 ACCGCGCACTTGTGCTGATTTGGGATGAGGCGGCTTACAGCTGAGAGTACCGCGCG 1140
1157 ACCGCGCACTTGTGCTGATTTGGGATGAGGCGGCTTACAGCTGAGAGTACCGCGCG 1216

677 TTCCCAAGGCGCAGTCTTCCAGCACCAGCACTGATCCAGGTTCCGGGTTTCT 736
661 AAGACTCTCAGTGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTTC 720
737 AAGACTCTCAGTGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTTC 796
721 CCCTCGCGCAAGCTCTCTAGAGCGGGGCGCGCGACCCCGCGCAGAGAGAGGCG 780
797 CCCTCGCGCAAGCTCTCTAGAGCGGGGCGCGCGACCCCGCGCAGAGAGAGGCG 856
781 GAGCGGGAGCGCGCGCGGAAAAGGCGCGCGAGGGGTCTCTGCAACCGCGCACTTGG 840
857 GAGCGGGAGCGCGCGCGGAAAAGGCGCGCGAGGGGTCTCTGCAACCGCGCACTTGG 916
841 CTTGCTCTCGTCCCGCGCGCACTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 900
917 CTTGCTCTCGTCCCGCGCGCACTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 976
901 CTTGCTCTCGTCCCGCGCGCACTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 960
977 CTTGCTCTCGTCCCGCGCGCACTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1036
961 TTGCGCCCCCGCGCGCAGAGAGCGGGAAGCGCGCGGTCCGCGCGAGGAGCTGCAG 1020
1037 TTGCGCCCCCGCGCGCAGAGAGCGGGAAGCGCGCGGTCCGCGCGAGGAGCTGCAG 1096
1021 TACTTGGGCGAGATCCCAACATCTCTGCTGGCGGTCTGAGAGAGCAACCGCACGGGC 1080
1097 TACTTGGGCGAGATCCCAACATCTCTGCTGGCGGTCTGAGAGAGCAACCGCACGGGC 1156
1081 ACCGCGCACTTGTGCTGATTTGGGATGAGGCGGCTTACAGCTGAGAGTACCGCGCG 1140
1157 ACCGCGCACTTGTGCTGATTTGGGATGAGGCGGCTTACAGCTGAGAGTACCGCGCG 1216
1141 GCGCCCTCGCGGAGCGGTTGGGGAAGAGGAGCGCGGTTGGGA 1187
1217 GCGCCCTCGCGGAGCGGTTGGGGAAGAGGAGCGCGGTTGGGA 1263

RESULT 3
US-09-880-107-1590
; Sequence 1590, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1590
; LENGTH: 18596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D00596
US-09-880-107-1590

Query Match 55.4%; Score 658; DB 10; Length 18596;
Best Local Similarity 99.2%; Pred. No. 4.2e-168;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
1 GATCGGCCACTGCACTCCAGCTTCAGGCTGGTGAGAGAGCGAGACTGTCTCAAAAAA 60

QY	1141	GGCCCCCTGGGACGGGTGGGGGAAAGGAGGAGCGGCGCTGGGGG	1187
Db	1217	GGCCCCCTGGGACGGGTGGGGGAAAGGAGGAGCGGCGCTGGGGG	1263
RESULT 4			
US-09-967-768A-119	; Sequence 119, Application US/09967768A		
US-09-967-768A-119	; Patent No. US20020150877A1		
US-09-967-768A-119	; GENERAL INFORMATION:		
US-09-967-768A-119	; APPLICANT: Augustus, Meena		
US-09-967-768A-119	; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu		
US-09-967-768A-119	; TITLE OF INVENTION: Sets		
US-09-967-768A-119	; FILE REFERENCE: 689290-72		
US-09-967-768A-119	; CURRENT APPLICATION NUMBER: US/09/967,768A		
US-09-967-768A-119	; CURRENT FILING DATE: 2001-09-28		
US-09-967-768A-119	; PRIOR APPLICATION NUMBER: US/60/236,109		
US-09-967-768A-119	; PRIOR FILING DATE: 2000-09-28		
US-09-967-768A-119	; PRIOR APPLICATION NUMBER: US/60/236,034		
US-09-967-768A-119	; PRIOR FILING DATE: 2000-09-28		
US-09-967-768A-119	; PRIOR APPLICATION NUMBER: US/60/236,111		
US-09-967-768A-119	; PRIOR FILING DATE: 2000-09-28		
US-09-967-768A-119	; NUMBER OF SEQ ID NOS: 325		
US-09-967-768A-119	; SOFTWARE: PatentIn version 3.0		
US-09-967-768A-119	; SEQ ID NO 119		
US-09-967-768A-119	; LENGTH: 18596		
US-09-967-768A-119	; TYPE: DNA		
US-09-967-768A-119	; ORGANISM: Homo sapiens		
US-09-967-768A-119	; US-09-967-768A-119		
QY	Query Match 55.4%; Score 658; DB 10; Length 18596;		
Db	Best Local Similarity 99.2%; Pred. No. 4.2e-168;		
QY	Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;		
Db	1 GATCGCGCACTGCACCTCCAGCTGGGTGAGAGCGGAGCTGTCTCAAAAAA 60		
QY	78 GATCGCGCACTGCACCTCCAGCTGGGTGAGAGCGGAGCTGTCTCAAAAAA 137		
Db	61 AAAAAAGCCGAGGCTCAACAAAAACCTCGGAAAGCCCTGGCGGTCTTTTTT 120		
QY	138 AAAAAAGCCGAGGCTCAACAAAAACCTCGGAAAGCCCTGGCGGTCTTTTTT 197		
Db	121 TT 180		
QY	198 TT 257		
Db	181 CGGATCTGGCTCACTGCACTCTGCCTCCAGGTTCAAGCAATTTCTTGCTCAGC 240		
QY	258 CGGATCTGGCTCACTGCACTCTGCCTCCAGGTTCAAGCAATTTCTTGCTCAGC 317		
Db	241 TCCCAAGTAGCCACCAACCCAGCTAATTTTTTTTANTTTTGTAGAGCGGGTTTCA 300		
QY	318 TCCCAAGTAGCCACCAACCCAGCTAATTTTTTTTANTTTTGTAGAGCGGGTTTCA 377		
Db	301 CATGTTGTCCAGGCTGGTCTNGAATCTCGAATCTCAGGTGATCCACCGGCTCGGCCCC 360		
QY	378 CATGTTGTCCAGGCTGGTCTNGAATCTCAGGTGATCCACCGGCTCGGCCCC 437		
Db	361 CAAAGTACTAGGATTTACAGGCTGAGCCCGGTCAGCGCCCTCGCGGTTTAAATCA 420		
QY	438 CAAAGTACTAGGATTTACAGGCTGAGCCCGGTCAGCGCCCTCGCGGTTTAAATCA 497		
Db	421 AGTAGAAAAGCTGCATATATACACTTCTGCTTCGTTTGTAGTACAGCAAGAAATGG 480		
QY	498 AGTAGAAAAGCTGCATATATACACTTCTGCTTCGTTTGTAGTACAGCAAGAAATGG 556		
Db	481 AAATGCAAAATCCTTATAGTTGTAGGAAACAGATCTCAAAACAGCAGTTTGTGACAAG 540		
QY	557 AAATGCAAAATCCTTATAGTTGTAGGAAACAGATCTCAAAACAGCAGTTTGTGACAAG 616		
Db	541 ACCCGAGGAAAGCTGGAACTGTCTCTGCTTTAGAGAAAGCGCGGTTCGACAGACGG 600		

Y 1082 CCGGCACTCTCGGATTCGGATCGAGCGCGCTACAGCTGAG 1129
 b 61 CCGGCACTCTCGGATTCGGATTCGGATCGAGCGCGCTACAGCTGAG 108

RESULT 6

US-09-764-869-2113/c
 ; Sequence 2113, Application US/09764869
 ; Patent No. US20020061521A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC007
 ; CURRENT APPLICATION NUMBER: US/09/764,869
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 2442
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2113
 ; LENGTH: 10867
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-764-869-2113

Query Match 4.9%; Score 58; DB 10; Length 10867;
 Best Local Similarity 100.0%; Pred. No. 2.8e-08;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 GCCACTGCACCTCCAGCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAA 64
 Db 5892 GCCACTGCACCTCCAGCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAA 5835

RESULT 7

US-09-764-877-3970/c
 ; Sequence 3970, Application US/09764877
 ; Patent No. US20020147140A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC005
 ; CURRENT APPLICATION NUMBER: US/09/764,877
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 4031
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3970
 ; LENGTH: 10867
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-764-877-3970

Query Match 4.9%; Score 58; DB 10; Length 10867;
 Best Local Similarity 100.0%; Pred. No. 2.8e-08;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 GCCACTGCACCTCCAGCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAA 64
 Db 5892 GCCACTGCACCTCCAGCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAA 5835

RESULT 8

US-09-764-869-2112/c
 ; Sequence 2112, Application US/09764869
 ; Patent No. US20020061521A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC007
 ; CURRENT APPLICATION NUMBER: US/09/764,869
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 2442
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2112
 ; LENGTH: 10907
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-764-869-2112

Query Match 4.9%; Score 58; DB 10; Length 10907;
 Best Local Similarity 100.0%; Pred. No. 2.8e-08;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 GCCACTGCACCTCCAGCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAA 64
 Db 5929 GCCACTGCACCTCCAGCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAA 5872

RESULT 9

US-09-764-877-3968/c
 ; Sequence 3968, Application US/09764877
 ; Patent No. US20020147140A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC005
 ; CURRENT APPLICATION NUMBER: US/09/764,877
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 4031
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3968
 ; LENGTH: 10907
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-764-877-3968

Query Match 4.9%; Score 58; DB 10; Length 10907;
 Best Local Similarity 100.0%; Pred. No. 2.8e-08;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 GCCACTGCACCTCCAGCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAA 64
 Db 5929 GCCACTGCACCTCCAGCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAA 5872

RESULT 10

US-09-764-877-3972/c
 ; Sequence 3972, Application US/09764877
 ; Patent No. US20020147140A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC005
 ; CURRENT APPLICATION NUMBER: US/09/764,877
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 4031
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3972
 ; LENGTH: 12718
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-764-877-3972

Query Match 4.9%; Score 58; DB 10; Length 12718;
 Best Local Similarity 100.0%; Pred. No. 2.6e-08;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 GCCACTGCACCTCCAGCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAA 64
 Db 5929 GCCACTGCACCTCCAGCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAA 5872

```
RESULT 11
US-09-729-920-3/c
; Sequence 3, Application US/09729920
; Patent No. US20020103115A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000858
; CURRENT APPLICATION NUMBER: US/09/729,920
; CURRENT FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 143306
; TYPE: DNA
; ORGANISM: Human
US-09-729-920-3

Query Match      4.7%; Score 56; DB 10; Length 143306;
Best Local Similarity 100.0%; Pred. No. 4.4e-08;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 GGCCTACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTTGGCTCAGCTCCC 244
|||||
Db 65791 GGCCTACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTTGGCTCAGCTCCC 65736

RESULT 12
US-09-764-869-2111/c
; Sequence 2111, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2111
; LENGTH: 10894
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-2111

Query Match      4.6%; Score 55; DB 10; Length 10894;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GCCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAAAAAA 61
|||||
Db 5911 GCCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAAAAAA 5857

RESULT 13
US-09-982-091A-5
; Sequence 5, Application US/09982091A
; Patent No. US20020151030A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: KUMAGAI, Akiko
; APPLICANT: DUNPHY, William
; TITLE OF INVENTION: CLASPIN PROTEINS AND METHODS OF USE THEREOF
; FILE REFERENCE: CIT1320-1
; CURRENT APPLICATION NUMBER: US/09/982,091A
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: US 60/241,246
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 5
; LENGTH: 58837
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-982-091A-5

Query Match      4.5%; Score 54; DB 10; Length 58837;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 TGCATCTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAAAAAA 65
|||||
Db 14914 TGCATCTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAAAAAA 14967

RESULT 14
US-09-764-860-1108
; Sequence 1108, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1108
; LENGTH: 32170
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-1108

Query Match      4.0%; Score 48; DB 10; Length 32170;
Best Local Similarity 100.0%; Pred. No. 9.2e-06;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 GCAACCTCTGCTCCAGGTTCAAGCAATTTCTTGGCTCAGCTCCC 244
|||||
Db 3592 GCAACCTCTGCTCCAGGTTCAAGCAATTTCTTGGCTCAGCTCCC 3639

RESULT 15
US-09-954-531-145/c
; Sequence 145, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 145
; LENGTH: 13500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(13500)
```

OTHER INFORMATION: n=a,t,g or c
JS-09-954-531-145

Query Match 3.9%; Score 46; DB 9; Length 13500;
Best Local Similarity 100.0%; Pred. No. 4e-05;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 184 ATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCTT 229
|||||
|||||

Db 9584 ATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCTT 9539
|||||
|||||

Search completed: November 29, 2002, 05:51:16
Job time : 584.685 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

DN nucleic - nucleic search, using sw model

Run on: November 29, 2002, 00:20:35 ; Search time 48.3867 Seconds
(without alignments)
7523.251 Million cell updates/sec

Title: US-09-963-333-7

Perfect score: 1187

Sequence: 1 gatcgccgactgcactcca.....gagggagcgcgctgggga 1187

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 15

Total number of hits satisfying chosen parameters: 23436

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*

5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	658	55.4	18596	4	US-09-318-448-11
2	297	25.0	1536	3	US-09-089-195-1
3	297	25.0	1536	4	US-09-367-007C-38
4	205	17.3	942	4	US-09-347-878-29
5	46	3.9	72604	4	US-09-268-992-7
6	46	3.9	72604	4	US-09-657-474-7
7	40	3.4	2900	3	US-09-038-832-1
8	40	3.4	4136	4	US-09-103-875-2
9	40	3.4	35100	5	PCT-US93-06251-19
10	40	3.4	35100	1	US-08-306-691B-19
11	39	3.3	1014	4	US-09-257-179-32
12	39	3.3	1386	2	US-08-687-080-76
13	39	3.3	5300	4	US-08-938-669A-1
14	39	3.3	5304	4	US-08-938-669A-2
15	39	3.3	6169	4	US-08-938-669A-3
16	39	3.3	72604	4	US-09-268-992-7
17	39	3.3	72604	4	US-09-657-474-7
18	39	3.3	112132	4	US-09-741-150-3
19	39	3.3	176373	3	US-09-128-155-17
20	38	3.2	38844	4	US-09-734-675-3
21	38	3.2	70000	4	US-09-851-896-3
22	38	3.2	162450	4	US-09-345-882-1
23	37	3.1	6792	4	US-09-374-454-20
24	37	3.1	43950	4	US-09-735-934A-3
25	36	3.0	461	4	US-09-404-879A-1
26	36	3.0	461	4	US-09-404-879A-3
27	36	3.0	589	4	US-09-385-982-455

ALIGNMENTS

RESULT 1

US-09-318-448-11

; Sequence 11, Application US/09318448

; Patent No. 6210350

; GENERAL INFORMATION:

; APPLICANT: Johnson, William G.

; APPLICANT: Stearros, Edward S.

; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING

; TITLE OF INVENTION: DEVELOPMENTAL DISORDERS

; FILE REFERENCE: 601-1-057

; CURRENT APPLICATION NUMBER: US/09/318,448

; CURRENT FILING DATE: 1999-05-25

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 11

; LENGTH: 18596

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-318-448-11

Query Match 55.4%; Score 658; DB 4; Length 18596;
Best Local Similarity 99.2%; Pred. No. 4.1e-215;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Qy	1	GATCGCGCCACTGCACCTCCAGCCTGGGTGAGAGAGCGAGACTCTGTCTCAAAAAAAAAA	60
Db	78	GATCGCGCCACTGCACCTCCAGCCTGGGTGAGAGAGCGAGACTCTGTCTCAAAAAAAAAA	137
Qy	61	AAAAGACGCCAGGCTCAACAAAACCTCGGAAAGCCCTGGGCTCTTTTTTTT	120
Db	138	AAAAGACGCCAGGCTCAACAAAACCTCGGAAAGCCCTGGGCTCTTTTTTTT	197
Qy	121	TTTTTTTTTTTTTTTTTTGGGACAGTCTGTCTGTGCGCAGGCTGGAGTACATCGT	180
Db	198	TTTTTTTTTTTTTTTTTTGGGACAGTCTGTCTGTGCGCAGGCTGGAGTACATCGT	257
Qy	181	CGGATCTTGGCTCACTGCAACCTCTCCCTCCAGGTTCAAGCAATTCCTTCCTCAGCC	240
Db	258	CGGATCTTGGCTCACTGCAACCTCTCCCTCCAGGTTCAAGCAATTCCTTCCTCAGCC	317
Qy	241	TCCCAAGTAGCACAACGCCAGCAATTTTGTGTTAGTAGAGCGGGTTTCAC	300
Db	318	TCCCAAGTAGCACAACGCCAGCAATTTTGTGTTAGTAGAGCGGGTTTCAC	377
Qy	301	CATGTTGTCAGGCTGTCTGACCTCTGACCTCAGGTATCCACCCCTCGGCCCCC	360
Db	378	CATGTTGTCAGGCTGTCTGACCTCTGACCTCAGGTATCCACCCCTCGGCCCCC	437
Qy	361	CBAAGTACTAGGATTACAGCGGTGAGCCACCGCGTCCAGCGCCTGGCGTTTAAATCA	420


```
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/US98/02145
; PRIOR FILING DATE: 1998-01-03
; NUMBER OF SEQ ID NOS: 39
; SEQ ID NO 38
; LENGTH: 1536
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Wild type human thymidylate synthase cDNA
; OTHER INFORMATION: (Genbank Accession number IM 001071)
; JS-09-367-007C-38

Query Match      25.0%; Score 297; DB 4; Length 1536;
Best Local Similarity 100.0%; Pred. No. 1.3e-92;
Matches 297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 833 CCACCTTGCCCTGCTCGTCCGCGCCGCGCCACTTGGCTTCCGTCGCGCGCCAC 892
Db 14 CCACCTTGCCCTGCTCGTCCGCGCCGCGCCACTTGGCTTCCGTCGCGCGCCAC 73

2Y 893 TTGCGCTGCTCGTCCGTCGCGCCGCGCCACTGCTGCGCGCTCGAGTGCCTG 952
Db 74 TTGCGCTGCTCGTCCGTCGCGCCGCGCCACTGCTGCGCGCTCGAGTGCCTG 133

2Y 953 GCGCGCTTGGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1012
Db 134 GCGCGCTTGGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 193

2Y 1013 AGCTGCACTACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1072
Db 194 AGCTGCACTACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 253

QY 1073 GCAGGCGACCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1129
Db 254 GCAGGCGACCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 310

RESULT 4
US-09-347-878-29
; Sequence 29, Application US/09347878C
; Patent No. 6376210
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
; FILE REFERENCE: 25885-1651
; CURRENT APPLICATION NUMBER: US/09/347,878C
; CURRENT FILING DATE: 1999-07-06
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Human thymidylate synthase gene: exons 1-8
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: D00596/GenBank
US-09-347-878-29

Query Match      17.3%; Score 205; DB 4; Length 942;
Best Local Similarity 100.0%; Pred. No. 2.1e-61;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 925 ATGCTGTGGCGGCTCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 984
Db 1 ATGCTGTGGCGGCTCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60

QY 985 GAGCGGACCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1044
Db 61 GAGCGGACCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120

QY 1045 CTCGCTGGCGGCTCAGGAAGGACGACCGCGCGCGCGCGCGCGCGCGCGCG 1104
```

```
Db 121 CTCGCTGCGCGCTCAGGAAGGACGACCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 1105 ATGCGGCGCGCTACAGCTGAGAG 1129
Db 181 ATGCGGCGCGCTACAGCTGAGAG 205

RESULT 5
US-09-268-992-7
; Sequence 7, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/268,992
; CURRENT FILING DATE: 1999-03-16
; EARLIER APPLICATION NUMBER: 09/236,134
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: 60/106,056
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/088,312
; EARLIER FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 60/078,044
; EARLIER FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-268-992-7

Query Match      3.9%; Score 46; DB 4; Length 72604;
Best Local Similarity 100.0%; Pred. No. 6.8e-08;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 GATCTTGCTCACTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 228
Db 46184 GATCTTGCTCACTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 46229

RESULT 6
US-09-657-474-7
; Sequence 7, Application US/09657474
; Patent No. 6399762
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/657,474
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/268,992
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/236,134
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/106,056
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/088,312
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/078,044
; PRIOR FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
```

; SEQ ID NO 7

; LENGTH: 72604

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: modified base

; LOCATION: all n positions

; OTHER INFORMATION: n=a, c, g, or t

; US-09-657-474-7

Query Match 3.9%; Score 46; DB 4; Length 72604;

Best Local Similarity 100.0%; Pred. No. 6.8e-08;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 GATCTGGCTCACTGCACTCTGCTCCAGGTTCAAGCAATTCT 228

Db 46184 GATCTGGCTCACTGCACTCTGCTCCAGGTTCAAGCAATTCT 46229

RESULT 7

US-09-038-832-1

; Sequence 1, Application US/09038832

; Patent No. 6146845

; GENERAL INFORMATION:

; APPLICANT: KIKLY, KRISTINE

; APPLICANT: BRICKSON-MILLER, CONNIE

; TITLE OF INVENTION: Sialoadhesin Family Member-2

; TITLE OF INVENTION: (SAF-2)

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RATNER & PRESTIA

; STREET: P.O. BOX 980

; CITY: VALLEY FORGE

; STATE: PA

; COUNTRY: USA

; ZIP: 19482

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/038,832

; FILING DATE: 11-MAR-1998

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/041,886

; FILING DATE: 02-APR-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: PRESTIA, PAUL F

; REGISTRATION NUMBER: 23,031

; REFERENCE/DOCKET NUMBER: GH-50018

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-407-0700

; TELEFAX: 610-407-0701

; TELEX: 846169

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2900 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; US-09-038-832-1

Query Match 3.4%; Score 40; DB 3; Length 2900;

Best Local Similarity 100.0%; Pred. No. 1.3e-05;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 GGCTCACTGCACTCTGCTCCAGGTTCAAGCAATTCT 228

Db 1881 GGCTCACTGCACTCTGCTCCAGGTTCAAGCAATTCT 1920

RESULT 8

US-09-103-875-2/c

; Sequence 2, Application US/09103875A

; Patent No. 6221849

; GENERAL INFORMATION:

; APPLICANT: Szyf, Moshe

; APPLICANT: Bigey, Pascal

; APPLICANT: Ramchandani, Shyam

; TITLE OF INVENTION: DNA METHYLTRANSFERASE GENOMIC SEQUENCES AND ANTISENSE

; TITLE OF INVENTION: OLIGONUCLEOTIDES

; FILE REFERENCE: 106101.194

; CURRENT APPLICATION NUMBER: US/09/103,875A

; CURRENT FILING DATE: 1998-06-24

; EARLIER APPLICATION NUMBER: 60/069,865

; EARLIER FILING DATE: 1997-12-17

; EARLIER APPLICATION NUMBER: 08/866,340

; EARLIER FILING DATE: 1997-05-30

; NUMBER OF SEQ ID NOS: 138

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 4136

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-09-103-875-2

Query Match 3.4%; Score 40; DB 4; Length 4136;

Best Local Similarity 100.0%; Pred. No. 1.3e-05;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 GATCTGGCTCACTGCACTCTGCTCCAGGTTCAAGC 222

Db 2510 GATCTGGCTCACTGCACTCTGCTCCAGGTTCAAGC 2471

RESULT 9

US-08-306-691B-19

; Sequence 19, Application US/08306691B

; Patent No. 5734039

; GENERAL INFORMATION:

; APPLICANT: Calabretta, Bruno

; APPLICANT: Skorski, Tomasz

; TITLE OF INVENTION: ANTISENSE

; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES

; NUMBER OF SEQUENCES: 55

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Seidel, Gonda, Lavoigna & Monaco, P.C.

; STREET: Two Penn Center, Suite 1800

; CITY: Philadelphia

; STATE: Pennsylvania

; COUNTRY: U.S.A.

; ZIP: 19102

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WordPerfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/306,691B

; FILING DATE: September 15, 1994

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Monaco, Daniel A.

; REGISTRATION NUMBER: 30,480

; REFERENCE/DOCKET NUMBER: 8321-8

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (215) 568-8383

; TELEFAX: (215) 568-5549

; TELEX: No. 5734039e

; INFORMATION FOR SEQ ID NO: 19:

```

; FILE REFERENCE: PZ015PI
; CURRENT APPLICATION NUMBER: US/09/257,179
; CURRENT FILING DATE: 1999-02-25
; EARLIER APPLICATION NUMBER: PCT/US98/17709
; EARLIER FILING DATE: 1998-08-27
; EARLIER APPLICATION NUMBER: 60/056,270
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,271
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,247
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,073
; EARLIER FILING DATE: 1997-08-29
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 32
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-257-179-32

Query Match          3.3%; Score 39; DB 4; Length 1014;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 GGTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 228
DB 940 GGTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 902

RESULT 12
US-08-687-080-76/c
; Sequence 76, Application US/08687080
; Patent No. 5965427
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,080
; FILING DATE: 17-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,126
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 76:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1386 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 35100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-306-691B-19

Query Match          3.4%; Score 40; DB 1; Length 35100;
Best Local Similarity 100.0%; Pred. No. 8.4e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 GGTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 228
DB 28051 GGTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 28090

RESULT 10
PCT-US93-06251-19
; Sequence 19, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06251
; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PCT-US93-06251-19

Query Match          3.4%; Score 40; DB 5; Length 35100;
Best Local Similarity 100.0%; Pred. No. 8.4e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 GGTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 228
DB 28051 GGTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 28090

RESULT 11
US-09-257-179-32/c
; Sequence 32, Application US/09257179
; Patent No. 6410709
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 29 Human Secreted Proteins

```

ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: INTRON 9 OF RAD50 GENOMIC SEQUENCE
US-08-687-080-76

Query Match 3.3%; Score 39; DB 2; Length 1386;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 ACTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGC 222
|||||
DB 586 ACTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGC 548

RESULT 13

US-08-938-669A-1
; Sequence 1, Application US/08938669A
; Patent No. 6171788
; GENERAL INFORMATION:
; APPLICANT: Nguyen, Thai D.
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS,
; TITLE OF INVENTION: PROGNOSIS AND TREATMENT OF GLAUCOMA AND
; TITLE OF INVENTION: RELATED DISEASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,669A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/791,154
; FILING DATE: 28-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mendelson, Elliot
; REGISTRATION NUMBER: P-42,878
; REFERENCE/DOCKET NUMBER: 07425-0034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 383-6857
; TELEFAX: 202 383-6610
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5300 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-938-669A-1

Query Match 3.3%; Score 39; DB 4; Length 5300;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 GCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 228
|||||
DB 1327 GCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 1365

RESULT 14

US-08-938-669A-2
; Sequence 2, Application US/08938669A
; Patent No. 6171788
; GENERAL INFORMATION:

; APPLICANT: Nguyen, Thai D.
; APPLICANT: Polansky, Jon R.
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS,
; TITLE OF INVENTION: PROGNOSIS AND TREATMENT OF GLAUCOMA AND
; TITLE OF INVENTION: RELATED DISEASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,669A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/791,154
; FILING DATE: 28-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mendelson, Elliot
; REGISTRATION NUMBER: P-42,878
; REFERENCE/DOCKET NUMBER: 07425-0034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 383-6857
; TELEFAX: 202 383-6610
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5304 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-938-669A-2

Query Match 3.3%; Score 39; DB 4; Length 5304;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 GCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 228
|||||
DB 1327 GCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 1365

RESULT 15

US-08-938-669A-3
; Sequence 3, Application US/08938669A
; Patent No. 6171788
; GENERAL INFORMATION:
; APPLICANT: Nguyen, Thai D.
; APPLICANT: Polansky, Jon R.
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS,
; TITLE OF INVENTION: PROGNOSIS AND TREATMENT OF GLAUCOMA AND
; TITLE OF INVENTION: RELATED DISEASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,669A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/791,154
; FILING DATE: 28-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mendelson, Elliot
; REGISTRATION NUMBER: P-42,878
; REFERENCE/DOCKET NUMBER: 07425-0034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 383-6857
; TELEFAX: 202 383-6610
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6169 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-938-669A-3
Query Match 3.3%; Score 39; DB 4; Length 6169;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 190 GCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 228
      ||||||||||||||||||||||||||||||||||
Db 1327 GCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 1365
```

Search completed: November 29, 2002, 03:34:35
Job time : 288.387 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

DM nucleic - nucleic search, using sw model

Run on: November 28, 2002, 22:40:30 ; Search time 235.395 Seconds
(without alignments)
11355.910 Million cell updates/sec

Title: US-09-963-333-7
Perfect score: 1187
Sequence: I gatcgccactgcactcca.....gagggggcgcggtgggga 1187

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 15

Total number of hits satisfying chosen parameters: 263288

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq 101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	DB ID	Description
1	55.4	18596	22	AAF31109		Thymidylate synthase
2	55.4	18596	22	AAC91215		Human thymidylate
3	55.4	18596	24	ABN95092		Gene #1590 used to
4	55.4	18596	24	ABK43334		Human Thymidylate
5	55.4	18596	24	ABL62854		Breast cancer rela
6	55.4	18596	24	ABL63078		Breast cancer relat
7	55.4	18596	24	ABL67927		Ovary cancer relat
8	297	25.0	1536	24	ABK43335	Human Thymidylate
9	262	22.1	3298	24	AAS94945	Human DNA sequence

C 10	205	17.3	1539	23	AAS94960	DNA encoding novel
C 11	165	13.9	1131	24	ABK43330	Human GNHK genomic
C 12	165	13.9	1161	24	ABK43288	Human cDNA encodin
C 13	108	9.1	535	24	ABL38559	Human colon tumour
C 14	62	5.2	7428	22	ABA07164	Human pancreatic c
C 15	60	5.1	566	20	AAK24270	Human thymidylate
C 16	59	5.0	22473	22	AAK71400	Human immune/haema
C 17	58	4.9	3702	22	AAK72556	Human immune/haema
C 18	58	4.9	10867	22	AAK37605	Human musculoskele
C 19	58	4.9	10867	22	AAK36613	Human cardiovascular
C 20	58	4.9	10867	22	AAK66345	Human immune/haema
C 21	58	4.9	10907	22	AAK37603	Human musculoskele
C 22	58	4.9	10907	22	AAK36612	Human cardiovascular
C 23	58	4.9	10907	22	AAK66344	Human immune/haema
C 24	58	4.9	12718	22	AAK37607	Human musculoskele
C 25	58	4.9	12718	22	AAK66346	Human immune/haema
C 26	58	4.9	23989	22	AAK72555	Human immune/haema
C 27	56	4.7	913	22	AAI94312	Human neuroblastom
C 28	56	4.7	5253	22	AAK9859	Human digestive sy
C 29	56	4.7	5257	22	AAK9860	Human digestive sy
C 30	56	4.7	143306	24	ABK49586	Human transporter
C 31	55	4.6	10894	22	AAK36611	Human cardiovascular
C 32	55	4.6	10894	22	AAK66343	Human immune/haema
C 33	54	4.5	10012	24	ABL55899	Human small induci
C 34	54	4.5	58837	24	ABK52612	Human Claspino geno
C 35	50	4.2	38358	22	AAK73555	Human immune/haema
C 36	48	4.0	32170	22	AAK28674	Genomic sequence #
C 37	48	4.0	36501	22	AAK64829	Human immune/haema
C 38	47	4.0	1110	19	AAV29343	Calcium ion channe
C 39	47	4.0	12278	22	ABA14585	Human nervous syst
C 40	47	4.0	12278	22	ABA14664	Human nervous syst
C 41	46	3.9	111	22	AAK06986	Human reproductive
C 42	46	3.9	115	22	AAK90875	Human digestive sy
C 43	46	3.9	255	21	AAC05034	Human secreted pro
C 44	46	3.9	429	22	AAK12261	Human breast cance
C 45	46	3.9	490	22	AAK77714	Human immune/haema

ALIGNMENTS

RESULT 1

AAF31109
ID AAF31109 standard; cDNA; 18596 BP.

XX AC AAF31109;

XX DT 27-APR-2001 (first entry)

XX DE Thymidylate synthase coding sequence.

XX DE Analyte-binding enzyme; analyte analysis; ss.

XX KW Homo sapiens.

XX OS WO200102600-A2.

XX PN 11-JAN-2001.

XX PD 30-JUN-2000; 2000WO-US18057.

XX PR 06-JUL-1999; 99US-0347878.

XX PR 06-DEC-1999; 99US-0457205.

XX XX (GEAT) GEN ATOMICS.

XX PI Yuan C;

XX DR WPI; 2001-071583/08.

XX PT Assaying method, useful for prognosis and diagnosis of disease,
PT comprises contacting sample with a mutant analyte-binding enzyme and
PT detecting binding -

XX Disclosure; Page -; 187pp; English.

XX The present invention relates to a method for assaying an analyte in a sample comprising: contacting the sample with a mutant analyte-binding enzyme which has binding affinity for the analyte or an immediate activity, and detecting resulting binding. The method is useful in monitoring biological systems/processes, or prognosis/diagnosis of disease caused by imbalances of the analytes. The present sequence is a coding sequence used in the present invention.

CC Note: the present sequence is not shown in the specification, but was from Genbank, using information given in the specification.

XX

SQ Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;

Query Match 55.4%; Score 658; DB 22; Length 18596;

Best Local Similarity 99.2%; Pred. No. 6.3e-195;

Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1 GATCGCGCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 60

Db 78 GATCGCGCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 137

QY 61 AAAAAGACCGCCAGGGCTCAAAACAAAACCTCGGAAGACCTCGGGTCTTTT 120

Db 138 AAAAGACCGCCAGGGCTCAAAACAAAACCTCGGAAGACCTCGGGTCTTTT 197

QY 121 TTTTCTTTTCTTTTGGGACAGTCTGTCTGTGCGCCAGGCTGGAGTACAATGGT 180

Db 198 TTTTCTTTTCTTTTGGGACAGTCTGTCTGTGCGCCAGGCTGGAGTACAATGGT 257

QY 181 CGGATCTGGCTCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAGC 240

Db 258 CGGATCTGGCTCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAGC 317

QY 241 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGTTTCA 300

Db 318 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGTTTCA 377

QY 301 CATGTGTCCAGGCTGGTCTGTAACCTCTGACCTCAGTGATCCACCGGCTCG 360

Db 378 CATGTGTCCAGGCTGGTCTGTAACCTCTGACCTCAGTGATCCACCGGCTCG 437

QY 361 CAAAGTACTAGGATTAACAGCGTGAGCCACCGGCTCAGCGCTTGGCGGTTTAA 420

Db 438 CAAAGTACTAGGATTAACAGCGTGAGCCACCGGCTCAGCGCTTGGCGGTTTAA 497

QY 421 AGTAGAAAGCTGCAATTAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480

Db 498 AGTAGAAAGCTGCAATTAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 556

QY 481 AAATGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540

Db 557 AAATGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 616

QY 541 ACAGCAAGAAACGTTGGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600

Db 617 ACAGCAAGAAACGTTGGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676

QY 601 TTTCCCAAGGCGGAGTCTTCCGCGCCAGCGCACTGCTGCTGCTGCTGCTGCTG 660

Db 677 TTTCCCAAGGCGGAGTCTTCCGCGCCAGCGCACTGCTGCTGCTGCTGCTGCTG 736

QY 661 AAGACTCTCAGCTGCGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 720

Db 737 AAGACTCTCAGCTGCGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 796

QY 721 CCCTTGGCGCAGCTCTTACAGCGGGGGCGCGGACCGCGCGGAGGAGAGGCG 780

Db 797 CCCTTGGCGCAGCTCTTACAGCGGGGGCGCGGACCGCGCGGAGGAGAGGCG 856

QY 781 GAGCGGGGACCGGCGCGGGAAGAGCGCGGGAAGGGGTCTGTGCCACCGCGCACTTGG 840

Db 857 GAGCGGGGACCGCGCGGGGAAAGGCGCGCGGAGGGGTCTCTGCCACCGCGCACTTGG 916

QY 841 CTTGCTCTCGTCCCGCGCGCCACATTGCGCTTCTCTCTCTCTCTCTCTCTCTCT 900

Db 917 CTTGCTCTCGTCCCGCGCGCCACATTGCGCTTCTCTCTCTCTCTCTCTCTCTCT 976

QY 901 CTTGCT 960

Db 977 CTTGCT 1036

QY 961 TTGCCCCCG 1020

Db 1037 TTGCCCCCG 1096

QY 1021 TACCTTGGGCGAGATCCAAACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080

Db 1097 TACCTTGGGCGAGATCCAAACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1156

QY 1081 ACGGCACT 1140

Db 1157 ACGGCACT 1216

QY 1141 GGCCT 1187

Db 1217 GGCCT 1263

RESULT 2

AAC91215

ID AAC91215 standard; DNA; 18596 BP.

XX AAC91215;

XX 20-MAR-2001 (first entry)

XX Human thymidylate synthase gene SEQ ID NO: 11.

XX Human; schizophrenia; developmental disorder; spina bifida cystica;

XX Tourette's syndrome; bipolar illness; autism; conduct disorder;

XX attention deficit disorder; obsessive compulsive disorder;

XX chronic multiple tic syndrome; learning disorder; polymorphism; ds.

XX Homo sapiens.

XX WC2000071754-A1.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000WO-US14354.

XX 25-MAY-1999; 99US-0318448.

XX (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.

XX Johnson WG, Stenroos ES;

XX WPI; 2001-025174/03.

XX

XX Diagnosing a developmental disorder, e.g. schizophrenia, by forming

XX datasets (DS) of genetic (e.g. genotypes of folate metabolism alleles)

XX and environmental variables affecting an individual and then comparing

XX these DS with reference DS -

XX Disclosure; Page 125-131; 156pp; English.

XX The present invention provides a novel method of estimating the

XX susceptibility of an individual to a developmental disorder using genetic

XX and environmental variables. The method can be used in the diagnosis,

XX prevention and treatment of disorders such as schizophrenia, spina bifida

XX cystica, Tourette's syndrome, bipolar illness, autism, conduct disorders,

XX attention deficit disorder, obsessive compulsive disorder, chronic

XX multiple tic syndrome and learning disorders such as dyslexia.

XX
SQ Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;
Query Match 55.4%; Score 658; DB 22; Length 18596;
Best Local Similarity 99.2%; Pred. No. 6.3e-195;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
1 GATCGCGCCACTGCACTCCAGCTCGGCTGAGAGAGCAGACTCTGTCTCAAAAAA 60
Db GATCGCGCCACTGCACTCCAGCTCGGCTGAGAGAGCAGACTCTGTCTCAAAAAA 137
61 AAAAGACGCGCAGGCTCAAAACAAAAAAGCTCGAAGAGCCTGCGGCTCTTTT 120
Db AAAAGACGCGCAGGCTCAAAACAAAAAAGCTCGAAGAGCCTGCGGCTCTTTT 197
121 TTTTCTTTTCTTTTCTTTTGGCAGCTCTGCTGCTGCGCCAGGCTGAGTACAAAGT 180
Db TTTTCTTTTCTTTTCTTTTGGCAGCTCTGCTGCTGCGCCAGGCTGAGTACAAAGT 257
198 TTTTCTTTTCTTTTCTTTTGGCAGCTCTGCTGCTGCGCCAGGCTGAGTACAAAGT 317
241 TCCCAAGTAGCCACGCGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGTTCCAC 300
Db TCCCAAGTAGCCACGCGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGTTCCAC 377
318 TCCCAAGTAGCCACGCGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGTTCCAC 497
361 CAAAGTACTAGGATTAAGCGGTGAGCAGCGCTCCAGCGCTCCAGCGCTTTTAAACA 420
Db CAAAGTACTAGGATTAAGCGGTGAGCAGCGCTCCAGCGCTCCAGCGCTTTTAAACA 497
421 AGTAGAAGAGCTGATTAACACTTGTCTTGTGTTGTTGTTGTTGTTGTTGTTGTT 480
Db AGTAGAAGAGCTGATTAACACTTGTCTTGTGTTGTTGTTGTTGTTGTTGTTGTT 556
481 AATGCAATCNCCTTATTTAGTTCTAGGAAACAGATCTCAAAACAGCAGTTTGTGACAG 540
Db AATGCAATCNCCTTATTTAGTTCTAGGAAACAGATCTCAAAACAGCAGTTTGTGACAG 616
541 ACCGAGAGAAACCTGGGAACTGCTGCTGCTTTAGAGAGCGCGGTGACACAGCG 600
Db ACCGAGAGAAACCTGGGAACTGCTGCTGCTTTAGAGAGCGCGGTGACACAGCG 676
617 ACCGAGAGAAACCTGGGAACTGCTGCTGCTTTAGAGAGCGCGGTGACACAGCG 780
Db ACCGAGAGAAACCTGGGAACTGCTGCTGCTTTAGAGAGCGCGGTGACACAGCG 856
601 TTCCCAAGAGGCGCAGTCTTCCCGCCACGCGCTGCTGCTGCTGCTGCTGCTGCT 736
Db TTCCCAAGAGGCGCAGTCTTCCCGCCACGCGCTGCTGCTGCTGCTGCTGCTGCT 796
661 AAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db AAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
737 AAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db AAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 856
721 CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
Db CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 856
797 CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
Db CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 916
781 GAGCGGGGAGCGCGCGGGGAAAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGG 900
Db GAGCGGGGAGCGCGCGGGGAAAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 976
841 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
917 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
901 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
977 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020

Db 1037 TTGCCCCCCCCCGCCACAGAGCGGAGCGCCGAGCCCGGTCGCCGACGCGGAGCTGCAG 1096
QY 1021 TACCTGGGCGAGATCAACACATCTCTCGCTGCGGCGTCAAGAGACGACCGACGCGGC 1080
Db 1097 TACCTGGGCGAGATCAACACATCTCTCGCTGCGGCGTCAAGAGACGACCGACGCGGC 1156
QY 1081 ACCGCGCACCTGTGCGGTATTCGCGATGACGCGCGCTTACGCTGAGAGTACCGCGC 1140
Db 1157 ACCGCGCACCTGTGCGGTATTCGCGATGACGCGCGCTTACGCTGAGAGTACCGCGC 1216
QY 1141 GCGCCCTGCGGAGCGGCTGCGGAGAGGAGGAGCGCGCTGGGA 1187
Db 1217 GCGCCCTGCGGAGCGGCTGCGGAGAGGAGGAGCGCGCTGGGA 1263
RESULT 3
ABN95092
ID ABN95092 standard; DNA; 18596 BP.
XX
AC ABN95092;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #1590 used to diagnose liver cancer.
XX
XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumor; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
XX WO200229103-A2.
XX
PD 11-APR-2002.
XX
XX 02-OCT-2001; 2001WO-US30589.
XX
XX 02-OCT-2000; 2000US-237054P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
XX WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX
XX Claim 1; SEQ ID NO 1590; 298pp; English.
XX
XX The invention relates to a novel method for diagnosing and detecting the
XX progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX tumor in a patient, and differentiating metastatic liver cancer from
XX hepatocellular carcinoma in a patient, involving detecting the level of
XX expression of two or more genes represented in ABN9503-ABN97455 in a
XX tissue sample. The method of the invention has hepatotropic, and
XX cytostatic activity. The method is useful for diagnosing and detecting
XX the progression of liver cancer, hepatocellular carcinoma and metastatic
XX liver carcinoma in a patient. The method is useful for identifying
XX expression profiles which serve as useful diagnostic markers as well as
XX markers that can be used to monitor disease states, disease progression,
XX drug toxicity, drug efficacy and drug metabolism.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;

Query Match 55.4%; Score 658; DB 24; Length 18596;
Best Local Similarity 99.2%; Pred. No. 6.3e-195;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1 GATCCGCCACTGCTCAGCTCGAGCTGGTGGAGAGCGAGACTCTCTCTCAAAAAA 60
Db 78 GATCGCGCCACTGCACTCCAGCTCGGTGAGAGAGCGAGACTCTCTCTCAAAAAA 137
QY 61 AAAAGACCGCCAGGCTCAACAAAAAACCCTCGAAAAAGCCCTGCGGTCTTTTTTT 120
Db 138 AAAAGACCGCCAGGCTCAACAAAAAACCCTCGAAAAAGCCCTGCGGTCTTTTTTT 197
QY 121 TTTTCTTTTCTTTTCTTTTGGGACAGTCTGCTCTGCGCCAGGCTGAGTACAAATG 180
Db 198 TTTTCTTTTCTTTTCTTTTGGGACAGTCTGCTCTGCGCCAGGCTGAGTACAAATG 257
QY 181 CGGATCTGGCTACTGCACTGCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 240
Db 258 CGGATCTGGCTACTGCACTGCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 317
QY 241 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTTGTAGTAGAGCGGGTTTCA 300
Db 318 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTTGTAGTAGAGCGGGTTTCA 377
QY 301 CATGTGTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 378 CATGTGTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 437
QY 361 CAAAGTACTAGGATTAAGCGGTGAGCAACCGGCTGAGCGCTGCTGCTGCTGCT 420
Db 438 CAAAGTACTAGGATTAAGCGGTGAGCAACCGGCTGAGCGCTGCTGCTGCTGCT 497
QY 421 AGTAGAAAGCTGATTAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 498 AGTAGAAAGCTGATTAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556
QY 481 AATGCAATCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 557 AATGCAATCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 616
QY 541 ACGCAGGAAAGCTGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 617 ACGCAGGAAAGCTGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676
QY 601 TTCCCAAGGCGCAGTCTTCCGCGCCACCGCAGCTGCTGCTGCTGCTGCTGCTG 660
Db 677 TTCCCAAGGCGCAGTCTTCCGCGCCACCGCAGCTGCTGCTGCTGCTGCTGCT 736
QY 661 AAGACTCTCAGTGTGGCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 737 AAGACTCTCAGTGTGGCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
QY 721 CCCTCGCGCAGCTCTTAGAGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
Db 797 CCCTCGCGCAGCTCTTAGAGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 856
QY 781 GAGCGGCGCAGCGCGCGGGAAGCGCGCGGAGGGGCTCTGCTGCTGCTGCTGCT 840
Db 857 GAGCGGCGCAGCGCGCGGGAAGCGCGGAGGGGCTCTGCTGCTGCTGCTGCT 916
QY 841 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 917 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
QY 901 CTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 977 CTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
QY 961 TTGCGCCCGCGCGCAGAGCGGGAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
Db 1037 TTGCGCCCGCGCGCAGAGCGGGAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1096
QY 1021 TACTGCGGCGAGTTCACACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1097 TACTGCGGCGAGTTCACACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156

QY 1081 ACGGCACCTGTCTGCTATTCGGCATGCGAGCGCGCTACGCTGAGAGTGACGCGCG 1140
Db 1157 ACGGCACCTGTCTGCTATTCGGCATGCGAGCGCGCTACGCTGAGAGTGACGCGCG 1216
QY 1141 GCGCCCTGCGGAGCGGCTGCGGGAAGGAGGAGGCGCGGCTGCGGGA 1187
Db 1217 GCGCCCTGCGGAGCGGCTGCGGGAAGGAGGAGGCGCGGCTGCGGGA 1263
RESULT 4
ABK43334
ID ABK43334 standard; DNA; 18596 BP.
XX
AC ABK43334;
XX
DT 05-JUN-2002 (first entry)
XX
DE Human Thymidylate synthase gene sequence.
XX
KW HKNG1; ds; gene; chromosome 18p; bipolar affective disorder; BAD;
KW severe bipolar affective (mood) disorder; BP-I; schizophrenia; TS;
KW Hong Kong new gene 1; antimanic; antidepressant; neuroleptic;
KW Thymidylate synthase.
XX
OS Homo sapiens.
XX
PW WO200210366-A2.
XX
PD 07-FEB-2002.
XX
XX 02-AUG-2001; 2001WO-US24417.
XX
PF 02-AUG-2000; 2000US-0631275.
XX
PR 28-NOV-2000; 2000US-0722544.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
PA (REGC) UNIV CALIFORNIA.
XX
XX Chen H, Freimer NB, Novak T;
XX WPI; 2002-195962/25.
XX
XX New nucleic acid molecule Hong Kong New Gene 1 (HKNG1), useful for
XX screening for molecules which modulate HKNG1 expression for the
XX treatment of bipolar disorder and schizophrenia -
XX Disclosure; Fig 44; 367pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule comprising a
XX nucleotide sequence that encodes a Hong Kong New Gene (HKNG) 1 gene
XX product. The human gene for HKNG1 is located on chromosome 18p in
XX an area associated with bipolar affective disorder, BAD. Also
XX included are an expression vector comprising the nucleic acid, a method
XX of host cell expressing the nucleic acid, and identifying an individual (at
XX risk of) having HKNG1-mediated disorder comprising detecting the presence
XX or absence of a polymorphism that correlates with an HKNG1 allele
XX associated with the disorder, where the presence of the polymorphism
XX indicates that the individual (is at risk of) having HKNG1-mediated
XX disorder. A (small molecule) compound which modulates (inhibits or
XX potentiates) expression of a HKNG1 gene or gene product in a human
XX individual is useful for the treatment of a HKNG1-mediated disorder
XX such as bipolar affective disorder (BAD), severe bipolar affective (mood)
XX disorder (BP-I) and schizophrenia. The present sequence is a
XX thymidylate synthase, TS, genomic DNA sequence. The gene for TS
XX overlaps that of HKNG1 and therefore TS may also be involved in the
XX diseases listed above.
XX
SQ Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;
Query Match 55.4%; Score 658; DB 24; Length 18596;
Best Local Similarity 99.2%; Pred. No. 6 3e-195;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

```
QY 1 GATCGCGCACTGCTCCAGCCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 60
Db 78 GATCGCGCACTGCTCCAGCCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 137
QY 61 AAAAAGACCGCCAGGCTCAAAAACCAAAACCTCGAAAAAGCCCTGCGGCTCTTTTTTT 120
Db 138 AAAAAGACCGCCAGGCTCAAAAACCAAAACCTCGAAAAAGCCCTGCGGCTCTTTTTTT 197
QY 121 TTTTCTTTTTTTTTTTTTTTGGAGACAGTCTGTCTCTGTCGCCAGCTCGAGTCAATGT 180
Db 198 TTTTCTTTTTTTTTTTTTTTGGAGACAGTCTGTCTCTGTCGCCAGCTCGAGTCAATGT 257
QY 181 CGGATCTTGCTGCTCACTGCACTCTGCTGCTCCAGGTTCAAGCAATTTCTTCGCTCAGCC 240
Db 258 CGGATCTTGCTGCTCACTGCACTCTGCTGCTCCAGGTTCAAGCAATTTCTTCGCTCAGCC 317
QY 241 TCCCAAGTAGCCACACCGCCCACTGCTTAATTTTTGTATNTTTTAGTAGAGCGGGTTTCAC 300
Db 318 TCCCAAGTAGCCACACCGCCCACTGCTTAATTTTTGTATNTTTTAGTAGAGCGGGTTTCAC 377
QY 301 CATGTTGTCAGGCTGCTTNGAACTCTGACCTCAGGTSATCCACCGGCTCGGCCCCC 360
Db 378 CATGTTGTCAGGCTGCTTNGAACTCTGACCTCAGGTSATCCACCGGCTCGGCCCCC 437
QY 361 CAAAGTACTAGATTACAGGCGTGAGCACCGGCTCCAGGCTCGGCGGTTTTTAATCA 420
Db 438 CAAAGTACTAGATTACAGGCGTGAGCACCGGCTCCAGGCTCGGCGGTTTTTAATCA 497
QY 421 AGTAGAAAAAGCTGCATTAATACCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 498 AGTAGAAAAAGCTGCATTAATACCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556
QY 481 AAATGCAAAATCTTTATTTAGTTGAGAAACAGATCTCAAACAGCAGTTTTTGTGACAG 540
Db 557 AAATGCAAAATCTTTATTTAGTTGAGAAACAGATCTCAAACAGCAGTTTTTGTGACAG 616
QY 541 ACCGCGAGAAAGCTGGAACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 600
Db 617 ACCGCGAGAAAGCTGGAACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 676
QY 601 TTCCCAAGGCGCAGTCTCTTCGCGCCACCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 677 TTCCCAAGGCGCAGTCTCTTCGCGCCACCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
QY 661 AAGACTCTCAGCTGTGGCTCTGGCTCTCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 737 AAGACTCTCAGCTGTGGCTCTGGCTCTCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
QY 721 CCCTTGGCGCAGCTCTCTAGAGCGGGCGCGCGCAGCCCGCGCGAGGAGAGAGCGG 780
Db 797 CCCTTGGCGCAGCTCTCTAGAGCGGGCGCGCGCAGCCCGCGCGAGGAGAGAGAGCGG 856
QY 781 GAGCGGGGAGCGCGCGGAGAAAGCGCGGAGAGAGGCTCTGCAACCGCGCCTCTTGG 840
Db 857 GAGCGGGGAGCGCGCGGAGAAAGCGCGGAGAGAGGCTCTGCAACCGCGCCTCTTGG 916
QY 841 CTGCTCTCGCTCCGCGCGCGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 917 CTGCTCTCGCTCCGCGCGCGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
QY 901 CTGCTCTCCGCGCGCGCGCGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 977 CTGCTCTCCGCGCGCGCGCGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1036
QY 961 TTGCGCGCGCGCGCAGAGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
Db 1037 TTGCGCGCGCGCGCAGAGAGCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1096
QY 1021 TACCTGGGCGAGTCCAAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1097 TACCTGGGCGAGATCCAAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1156
```

QY 1081 ACCGGCACCCCTGCTGGTATTCGGCATGCAAGGCGGCTACAGCCTGAGAGTCAAGCCGCG 1140
Db 1157 ACCGGCACCCCTGCTGGTATTCGGCATGCAAGGCGGCTACAGCCTGAGAGTCAAGCCGCG 1216
QY 1141 GGCCCCCTGGGAGCGGCTGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1187
Db 1217 GGCCCCCTGGGAGCGGCTGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1263

RESULT 5
ABL62854
ID ABL62854 standard; DNA; 18596 BP.
XX ABL62854;
AC
XX
DT 15-MAY-2002 (first entry)
XX
DE Breast cancer related gene sequence SEQ ID NO:1191.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233172P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.

Db 1097 TACCTGGGAGATCCACACATCTCCGCTGCGGCTCAGGAAGACGACCGACGGC 1156
Qy 1081 ACCGGCACCTGTCGTTATTCGGCATGTCAGCGCGCTTACAGGCTGAGAGGTGACCGCGG 1140
Db 1157 ACCGGCACCTGTCGTTATTCGGCATGTCAGCGCGCTTACAGGCTGAGAGGTGACCGCGG 1216
Qy 1141 GCGCCCTGCGGACGGTGGCGGAGGAGGAGGCGCGCTGGGA 1187
Db 1217 GCGCCCTGCGGACGGTGGCGGAGGAGGAGGCGCGCTGGGA 1263

RESULT 7

ABL67927
ID ABL67927 standard; DNA; 18596 BP.

XX AC ABL67927;

XX DT 15-MAY-2002 (first entry)

XX DT Ovary cancer related gene sequence SEQ ID NO:6264.

XX KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
XX KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX KW gene; ds.

XX OS Homo sapiens.

XX XX WO200194629-A2.

XX XX 13-DEC-2001.

XX XX 30-MAY-2001; 2001WO-US10838.

XX XX 05-JUN-2000; 2000US-209473P.

XX XX 05-JUN-2000; 2000US-209531P.

XX XX 18-SEP-2000; 2000US-233133P.

XX XX 18-SEP-2000; 2000US-233617P.

XX XX 20-SEP-2000; 2000US-234009P.

XX XX 20-SEP-2000; 2000US-234034P.

XX XX 20-SEP-2000; 2000US-234052P.

XX XX 22-SEP-2000; 2000US-234509P.

XX XX 22-SEP-2000; 2000US-234567P.

XX XX 25-SEP-2000; 2000US-234923P.

XX XX 25-SEP-2000; 2000US-234924P.

XX XX 25-SEP-2000; 2000US-235077P.

XX XX 25-SEP-2000; 2000US-235082P.

XX XX 25-SEP-2000; 2000US-235134P.

XX XX 25-SEP-2000; 2000US-235280P.

XX XX 26-SEP-2000; 2000US-235637P.

XX XX 26-SEP-2000; 2000US-235638P.

XX XX 27-SEP-2000; 2000US-235711P.

XX XX 27-SEP-2000; 2000US-235720P.

XX XX 27-SEP-2000; 2000US-235840P.

XX XX 27-SEP-2000; 2000US-235863P.

XX XX 28-SEP-2000; 2000US-236028P.

XX XX 28-SEP-2000; 2000US-236032P.

XX XX 28-SEP-2000; 2000US-236033P.

XX XX 28-SEP-2000; 2000US-236034P.

XX XX 28-SEP-2000; 2000US-236109P.

XX XX 28-SEP-2000; 2000US-236111P.

XX XX 29-SEP-2000; 2000US-236842P.

XX XX 29-SEP-2000; 2000US-236891P.

XX XX 02-OCT-2000; 2000US-237172P.

XX XX 02-OCT-2000; 2000US-237173P.

XX XX 02-OCT-2000; 2000US-237278P.

XX XX 02-OCT-2000; 2000US-237294P.

XX XX 02-OCT-2000; 2000US-237295P.

XX XX 02-OCT-2000; 2000US-237316P.

XX XX 03-OCT-2000; 2000US-237598P.

XX XX 03-OCT-2000; 2000US-237604P.

PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX XX (AVAL-) AVALON PHARM.
XX XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX XX WPI; 2002-188264/24.
XX XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX XX
XX XX Claim 1; SEQ ID 6264; 44pp; English.
XX XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX XX
SQ Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;

Query Match 55.4%; Score 658; DB 24; Length 18596;

Best Local Similarity 99.2%; Pred. No. 6.3e-195;

Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Qy 1 GATCGCGCACTGCATCTCCAGCTGGGTGAGAGCGAGAGACTCTCTCAAAAAA 60

Db 78 GATCGCGCACTGCATCTCCAGCTGGGTGAGAGCGAGAGACTCTCTCAAAAAA 137

Qy 61 AAAAAAGCGCGAGCGCTCAACAAAAAACCTCGGAAAGCCCTGGGGCTTTTTTTT 120

Db 138 AAAAAAGCGCGAGCGCTCAACAAAAAACCTCGGAAAGCCCTGGGGCTTTTTTTT 197

Qy 121 TTTTTTTTTTTTTTTTTTTTGGGACAGTCTTGCTGCGCCAGGCTGGAGTACATG 180

Db 198 TTTTTTTTTTTTTTTTTTTTGGGACAGTCTTGCTGCGCCAGGCTGGAGTACATG 257

Qy 181 CGGATCTTGGCTCACTCAACCTCTCCCTCCAGGTTCAAGCAATTTCTGCTCAGCC 240

Db 258 CGGATCTTGGCTCACTCAACCTCTCCCTCCAGGTTCAAGCAATTTCTGCTCAGCC 317

Qy 241 TCCCAAGTAGCCACACGCCAGCTAATTTTGTATTTTGTATTTAGTAGAGCGGGGTTT 300

Db 318 TCCCAAGTAGCCACACGCCAGCTAATTTTGTATTTTGTATTTAGTAGAGCGGGGTTT 377

Qy 301 CATGTTGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360

Db 378 CATGTTGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 437

Qy 361 CAAAGTACTAGGATTACAGCGGAGCGACCGCTCCAGGCTGCTGCTGCTGCTGCTGCTGCT 420

Db 438 CAAAGTACTAGGATTACAGCGGAGCGACCGCTCCAGGCTGCTGCTGCTGCTGCTGCTGCT 497

Qy 421 AGTAGAAAAAGCTGCATTATACCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

Db 498 AGTAGAAAAAGCTGCATTATACCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556

```
QY 481 AAATGCAATCNCCTTATTAGTTGTAGGAAAACAGATCTCAAAACAGCAGTTTGTGACAAAG 540
Ddb 557 AAATGCAATCNCCTTATTAGTTGTAGGAAAACAGATCTCAAAACAGCAGTTTGTGACAAAG 616
QY 541 ACCGAGGAAACAGTGGGAACTGTCTCTCTGGCTTAGAGAAAGCGCGGTTCAGACAGACGG 600
Ddb 617 ACCGAGGAAACAGTGGGAACTGTCTCTCTGGCTTAGAGAAAGCGCGGTTCAGACAGACGG 676
QY 601 TTCCCAAGAGGCGAGTCTCTCCNCGCCACCGCACTGTGNTCCAGTTTCCCGGGTNTCCT 660
Ddb 677 TTCCCAAGAGGCGAGTCTCTCCNCGCCACCGCACTGTGNTCCAGTTTCCCGGGTNTCCT 736
QY 661 AAGACTCTACGTGTGGGCTCTGGGCTCTGGTCTGTGTCACACCCGTGGTCTCTGGTTC 720
Ddb 737 AAGACTCTACGTGTGGGCTCTGGGCTCTGGTCTGTGTCACACCCGTGGTCTCTGGTTC 796
QY 721 CCCCTGGCGACGCTCTCTAGAGCGGGGCGCGGACCCCGCGAGGAGAGAGGCG 780
Ddb 797 CCCCTGGCGACGCTCTCTAGAGCGGGGCGCGGACCCCGCGAGGAGAGAGGCG 856
QY 781 GAGCGGAGGACGCGCGGAGAAAGCGCGGAGAGGGGTCTGTGCAACCGCGCCACTTGG 840
Ddb 857 GAGCGGAGGACGCGCGGAGAAAGCGCGGAGAGGGGTCTGTGCAACCGCGCCACTTGG 916
QY 841 CCTGCTCTCTCCCGCGCGCCACTTGTGCTCTGCTCTCCGTCGCGCCACTTGGCTTG 900
Ddb 917 CCTGCTCTCTCCCGCGCGCCACTTGTGCTCTGCTCTCCGTCGCGCCACTTGGCTTG 976
QY 901 CCTCGCTCTCCCGCGCGCCACTTGTGCTCTGCTCTCCGTCGCGCCACTTGGCTTG 960
Ddb 977 CCTCGCTCTCCCGCGCGCCACTTGTGCTCTGCTCTCCGTCGCGCCACTTGGCTTG 1036
QY 961 TTGCCCCCGCGCGACAGAGCGGACCGCGAGCGCGCTTCCGCGCGAGGAGTGCAG 1020
Ddb 1037 TTGCCCCCGCGCGACAGAGCGGACCGCGAGCGCGCTTCCGCGCGAGGAGTGCAG 1096
QY 1021 TACCTGGGCGAGATCAACACATCTCTCGCTCTGCGCTCTGAGGAGGACCGACCGCGG 1080
Ddb 1097 TACCTGGGCGAGATCAACACATCTCTCGCTCTGCGCTCTGAGGAGGACCGACCGCGG 1156
QY 1081 ACCGCGACCTCTGTGCTGTTTGGCATGAGCGCGGCTACAGCTCTGAGAGGTGACCGCGG 1140
Ddb 1157 ACCGCGACCTCTGTGCTGTTTGGCATGAGCGCGGCTACAGCTCTGAGAGGTGACCGCGG 1216
QY 1141 GCGCCCTCTGCGGACGCGTCTGCGGAGAGGAGGAGGCGCGCTGCGGA 1187
Ddb 1217 GCGCCCTCTGCGGACGCGTCTGCGGAGAGGAGGAGGCGCGCTGCGGA 1263
RESULT 8
ABK43335
ID ABK43335 standard; cDNA; 1536 BP.
XX
XX ABK43335;
XX AC
XX DT
XX 05-JUN-2002 (first entry)
XX
DE Human Thymidylate synthase cDNA sequence.
XX
XX HKNG1; ss; gene; chromosome 18p; bipolar affective disorder; BAD;
KW severe bipolar affective (mood) disorder; BP-I; schizophrenia; TS;
KW Hong Kong new gene 1; antinomic; antidepressant; neuroleptic;
KW Thymidylate synthase.
XX
XX Homo sapiens.
XX OS
XX WO200210366-A2.
XX PN
XX 07-FEB-2002.
XX PD
XX 02-AUG-2001; 2001WO-US24417.
XX PP
XX DE Human DNA sequence #200 expressed during foam cell differentiation.
```

```
PR 02-AUG-2000; 2000US-0631275.
PR 28-NOV-2000; 2000US-0722544.
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA (REGC ) UNIV CALIFORNIA.
XX
XX Chen H, Freimer NB, Novak T;
PI
PI WPI; 2002-195962/25.
XX
XX P-PSDB; AAU87071.
XX
XX New nucleic acid molecule Hong Kong New Gene 1 (HKNG1), useful for
PT screening for molecules which modulate HKNG1 expression for the
PT treatment of bipolar disorder and schizophrenia -
XX
XX Disclosure; Fig 45; 367pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule comprising a
CC nucleotide sequence that encodes a Hong Kong New Gene (HKNG1) 1 gene
CC product. The human gene for HKNG1 is located on chromosome 18p in
CC an area associated with bipolar affective disorder. BAD. Also
CC included are an expression vector comprising the nucleic acid, a
CC host cell expressing the nucleic acid, an anti-HKNG1 antibody, a method
CC of identifying modulators of HKNG1, and identifying an individual (at
CC risk of) having HKNG1-mediated disorder comprising detecting the presence
CC or absence of a polymorphism that correlates with an HKNG1 allele
CC associated with the disorder, where the presence of the polymorphism
CC indicates that the individual (is at risk of) having HKNG1-mediated
CC disorder. A (small molecule) compound which modulates (inhibits or
CC potentiates) expression of a HKNG1 gene or gene product in a human
CC individual is useful for the treatment of a HKNG1-mediated disorder
CC such as bipolar affective disorder (BAD), severe bipolar affective (mood)
CC disorder (BP-I) and schizophrenia. The present sequence is the cDNA
CC encoding thymidylate synthase, TS. The gene for TS
CC overlaps that of HKNG1 and therefore TS may also be involved in the
CC diseases listed above.
XX
XX Sequence 1536 BP; 390 A; 369 C; 399 G; 378 T; 0 other;
SQ
Query Match 25.0%; Score 297; DB 24; Length 1536;
Best Local Similarity 100.0%; Pred. No. 3.4e-83;
Matches 297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 833 CCACCTTGGCTCTGCTCTCCGTCGCGCGCCACTTGGCTCTGCTCTCCGTCGCGCGCCAC 892
Ddb 14 CCACCTTGGCTCTGCTCTCCGTCGCGCGCCACTTGGCTCTGCTCTCCGTCGCGCGCCAC 73
QY 893 TTGCGCTCTGCTCTCCGTCGCGCGCCACTTGGCTCTGCTCTGCGCGCTCGAGCTGCGCG 952
Ddb 74 TTGCGCTCTGCTCTCCGTCGCGCGCCACTTGGCTCTGCTCTGCGCGCTCGAGCTGCGCG 133
QY 953 GCGCGCTCTGCTCTCCGTCGCGCGCCACTTGGCTCTGCGCGCTCGAGCTGCGCGG 1012
Ddb 134 GCGCGCTCTGCTCTCCGTCGCGCGCCACTTGGCTCTGCGCGCTCGAGCTGCGCGG 193
QY 1013 AGCTGAGTACTGCGGCGAGATCCACATCTCTGCTGCGCGCTCAGGAGGAGGAGGAC 1072
Ddb 194 AGCTGAGTACTGCGGCGAGATCCACATCTCTGCTGCGCGCTCAGGAGGAGGAGGAC 253
QY 1073 GACGCGGACCGGACCTCTGCTGTTATTCGGCATGACGCGCGCTACAGCTGAGAG 1129
Ddb 254 GCACGCGGACCGGACCTCTGCTGTTATTCGGCATGACGCGCGCTACAGCTGAGAG 310
RESULT 9
AAS94945
ID AAS94945 standard; DNA; 3298 BP.
XX
XX AAS94945;
XX AC
XX 14-FEB-2002 (first entry)
XX DT
XX Human DNA sequence #200 expressed during foam cell differentiation.
```



```
QY 1105 ATGCAGGCGGCTACAGCTGAGAG 1129
|||||
Db 1359 ATGCAGGCGGCTACAGCTGAGAG 1335
|||||

RESULT 11
ABK43330/c
ID ABK43330 standard; DNA; 1131 BP.
AC ABK43330;
XX
DT 05-JUN-2002 (first entry)
XX
DE Human GNKH genomic DNA.
XX
KW HKNG1; ds; gene; chromosome 18p; bipolar affective disorder; BAD;
KW severe bipolar affective (mood) disorder; BP-I; schizophrenia;
KW Hong Kong new gene 1; antimanic; antidepressant; neuroleptic.
XX
OS Homo sapiens.
XX WO200210366-A2.
XX
PD 07-FEB-2002.
XX
PF 02-AUG-2001; 2001WO-US24417.
XX
PR 02-AUG-2000; 2000US-0631275.
PR 28-NOV-2000; 2000US-0722544.
XX
PA (MILL-) MILLENNIUM PHARM INC.
PA (REGC ) UNIV CALIFORNIA.
PI Chen H, Freimer NB, Novak T;
XX
WPI; 2002-195962/25.
XX
DR New nucleic acid molecule Hong Kong New Gene 1 (HKNG1), useful for
PT screening for molecules which modulate HKNG1 expression for the
PT treatment of bipolar disorder and schizophrenia -
PS Example 16; Fig 30; 367pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule comprising a
CC nucleotide sequence that encodes a Hong Kong New Gene (HKNG) 1 gene
CC product. The human gene for HKNG1 is located on chromosome 18p in
CC an area associated with bipolar affective disorder, BAD. Also
CC included are an expression vector comprising the nucleic acid, a
CC host cell expressing the nucleic acid, an anti-HKNG1 antibody, a method
CC of identifying modulators of HKNG1, and identifying an individual (at
CC risk of) having HKNG1-mediated disorder comprising detecting the presence
CC or absence of a polymorphism that correlates with an HKNG1 allele
CC associated with the disorder, where the presence of the polymorphism
CC indicates that the individual (is at risk of) having HKNG1-mediated
CC disorder. A (small molecule) compound which modulates (inhibits or
CC potentiates) expression of a HKNG1 gene or gene product in a human
CC individual is useful for the treatment of a HKNG1-mediated disorder
CC such as bipolar affective disorder (BAD), severe bipolar affective (mood)
CC disorder (BP-I) and schizophrenia. The present sequence is HKNG1
CC genomic DNA sequence.
XX
SQ Sequence 1131 BP; 255 A; 316 C; 304 G; 254 T; 2 other;

Query Match 13.9%; Score 165; DB 24; Length 1131;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1023 CTGGGGGAGATCCACATCTCCGCTGGCGGTACGAGAGACGACCGGCGAC 1082
|||||
Db 789 CTGGGGGAGATCCACATCTCCGCTGGCGGTACGAGAGACGACCGGCGAC 730
|||||
QY 1083 CGGACCCCTGTGGTATTTCGGATGCGGCGCTACAGCTGAGAGTTGACCGCGGG 1142
|||||

Db 729 CGGACCCCTGTGGTATTTCGGATGCGGCGCTACAGCTGAGAGTTGACCGCGGG 670
QY 1143 CCCCTGCGGAGCGGTTGCGCGGAAGAGAGGAGCGCGCTGGGGA 1187
|||||
Db 669 CCCCTGCGGAGCGGTTGCGCGGAAGAGAGGAGCGCGCTGGGGA 625

RESULT 12
ABK43288/c
ID ABK43288 standard; cDNA; 1161 BP.
XX
AC ABK43288;
XX
DT 05-JUN-2002 (first entry)
XX
DE Human cDNA encoding GNKH protein.
XX
KW HKNG1; chromosome 18p; bipolar affective disorder; BAD; GNKH1; ss;
KW severe bipolar affective (mood) disorder; BP-I; schizophrenia; gene;
KW Hong Kong new gene 1; antimanic; antidepressant; neuroleptic.
XX
OS Homo sapiens.
XX WO200210366-A2.
XX
PD 07-FEB-2002.
XX
PF 02-AUG-2001; 2001WO-US24417.
XX
PR 02-AUG-2000; 2000US-0631275.
PR 28-NOV-2000; 2000US-0722544.
XX
PA (MILL-) MILLENNIUM PHARM INC.
PA (REGC ) UNIV CALIFORNIA.
PI Chen H, Freimer NB, Novak T;
XX
WPI; 2002-195962/25.
XX
DR New nucleic acid molecule Hong Kong New Gene 1 (HKNG1), useful for
PT screening for molecules which modulate HKNG1 expression for the
PT treatment of bipolar disorder and schizophrenia -
PS Example 16; Fig 28; 367pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule comprising a
CC nucleotide sequence that encodes a Hong Kong New Gene (HKNG) 1 gene
CC product. The human gene for HKNG1 is located on chromosome 18p in
CC an area associated with bipolar affective disorder, BAD. Also
CC included are an expression vector comprising the nucleic acid, a
CC host cell expressing the nucleic acid, an anti-HKNG1 antibody, a method
CC of identifying modulators of HKNG1, and identifying an individual (at
CC risk of) having HKNG1-mediated disorder comprising detecting the presence
CC or absence of a polymorphism that correlates with an HKNG1 allele
CC associated with the disorder, where the presence of the polymorphism
CC indicates that the individual (is at risk of) having HKNG1-mediated
CC disorder. A (small molecule) compound which modulates (inhibits or
CC potentiates) expression of a HKNG1 gene or gene product in a human
CC individual is useful for the treatment of a HKNG1-mediated disorder
CC such as bipolar affective disorder (BAD), severe bipolar affective (mood)
CC disorder (BP-I) and schizophrenia. The present sequence is a cDNA
CC encoding the GNKH1 protein, a protein which is encoded on the opposite
CC strand of DNA encoding HKNG1 and may also be involved in the diseases
CC listed above.
XX
SQ Sequence 1161 BP; 285 A; 316 C; 304 G; 254 T; 2 other;

Query Match 13.9%; Score 165; DB 24; Length 1161;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1023 CCTGGGGGAGATCCACATCTCCGCTGGCGGTACGAGAGACGACCGGCGAC 1082
|||||
```


Db	789	CTCGGGCAGATCCACACATCTCTCCGTCGGCGGTTCAGGAAGACACACCGCACGGGCAC	730
Qy	1083	CGGCACCCCTGTTCGGTATTCGGCATGCAGGCGCGCTACAGCTGAGAGTGAACCCCGCGG	1142
Db	729	CGGCACCCCTGTTCGGTATTCGGCATGCAGGCGCGCTACAGCTGAGAGTGAACCCCGCGG	670
Qy	1143	CCCTCGCGGACCGGCTGGCGGGAAGGAGGAGGCGCGCTGGGGGA	1187
Db	669	CCCTCGCGGACCGGCTGGCGGGAAGGAGGAGGCGCGCTGGGGGA	625
RESULT 13			
ID	ABL38559	standard; cDNA; 535 BP.	
AC	ABL38559;		
XX	08-APR-2002	(first entry)	
DE	Human colon tumour antigen polynucleotide SEQ ID NO:2148.		
XX	Human; colon cancer; colon tumour antigen; cytostatic; vaccine;		
KW	colon tumour metastatic antigen; diagnosis; gene; ss.		
XX	Homo sapiens.		
OS	WO200196388-A2.		
PN	20-DEC-2001.		
PD	08-JUN-2001; 2001WO-US18557.		
PF	09-JUN-2000; 2000US-210899P.		
PR	20-FEB-2001; 2001US-270216P.		
XX	(CORI-) CORIXA CORP.		
PA	Jiang Y, Harlocker SL, Secretist H;		
PI	WPI; 2002-114514/15.		
XX	Novel isolated colon tumor polynucleotide differentially expressed in		
PT	colon tumor or colon metastatic tumor and polypeptides encoded by them,		
PT	useful for inhibiting development of cancer in patient -		
XX	Claim 1; SEQ ID 2148; 105pp; English.		
XX	ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)		
CC	which were isolated from human colon tumour and colon metastatic tumour		
CC	cDNA libraries. (I) have cytostatic activity and can be used in vaccine		
CC	production. (I) can be used for stimulating and/or expanding T cells		
CC	specific for a tumour protein on contact with the T cells. They are also		
CC	useful for inhibiting the development of cancer in a patient. (I) can be		
CC	used as probes or primers for nucleic acid hybridisation, for preparing		
CC	mutant species primers, or primers for use in genetic constructions. (I)		
CC	can be used in the diagnosis of a colon tumour.		
XX	Sequence 535 BP; 135 A; 129 C; 151 G; 118 T; 2 other;		
SQL	Query Match	9.1%; Score 108; DB 24; Length 535;	
	Best Local Similarity	100.0%; Pred. No. 1e-24;	
	Matches 108; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1022	ACCTGGGCGAGATCCACACATCTCTCCGTCGGCGGTTCAGGAAGGACGACGACGCGCA	1081
Db	1	ACCTGGGCGAGATCCACACATCTCTCCGTCGGCGGTTCAGGAAGGACGACGCGCGCA	60
Qy	1082	CCGGCACCCCTGTTCGGTATTCGGCATGCAGGCGCGCTACAGCTGAGAG	1129
Db	61	CCGGCACCCCTGTTCGGTATTCGGCATGCAGGCGCGCTACAGCTGAGAG	108
RESULT 14			

```
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0232404.
PR 14-SEP-2000; 2000US-0232406.
PR 14-SEP-2000; 2000US-0232407.
PR 21-SEP-2000; 2000US-0232423.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0234984.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246529.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.

PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-457717/49.
XX
PT Isolated pancreatic cancer polypeptide for treating, preventing and/ or
PT prognosing disorders related to the pancreas including pancreatic
PT cancers and also for testing and detection e.g. diagnosis -
XX
PS Disclosure; SEQ ID NO 483; 537pp; English.
XX
CC The invention relates to an isolated polypeptide comprising an amino
CC acid sequence at least 90% identical to 188 amino acid sequences fully
CC defined in the specification and encoded by 188 cDNA clones fully
CC defined in the specification. The invention also relates to a fragment
CC having biological activity, a domain, an epitope, full length protein,
CC variant, allelic variant or a species homologue of the fully defined
CC sequence. The polynucleotide and polypeptide are useful for treating,
CC preventing and/or prognosing disorders related to the pancreas including
CC pancreatic cancer, pancreatitis, diabetes, endocrine disorders such as
CC acromegaly or hyperthyroidism, and gastrointestinal disorders such as
CC Crohn's disease and duodenal ulcers. The present sequence encodes a
CC pancreatic cancer-related polypeptide of the invention.
XX
SQ Sequence 7428 BP; 2045 A; 1352 C; 1528 G; 2503 T; 0 other;

Query Match 5.2%; Score 62; DB 22; Length 7428;
Best Local Similarity 100.0%; Pred. No. 8.3e-11;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 GATCTTGCTCACTGCAACCTGCTCCAGGTTCAAGCAATTTCTGCTCAGCCTC 242
Db 1692 GATCTTGCTCACTGCAACCTGCTCCAGGTTCAAGCAATTTCTGCTCAGCCTC 1751

QY 243 CC 244
Db 1752 CC 1753

RESULT 15
AAX24270
ID AAX24270 standard; DNA; 566 BP.
XX
AC AAX24270;
XX
DT 02-JUL-1999 (first entry)
XX
DE Human thymidylate synthase DNA.
XX
KW Thymidylate synthase; antisense oligonucleotide; cytostatic;
KW tumour cell; anticancer drug; treatment; cancer; antiproliferative;
KW antioestrogen; progesterone; antiandrogen; testosterone inhibitor;
KW anti-invasion agent; growth factor inhibitor; antimetabolite;
KW antibiotic; alkylating agent; antimitotic agent; radiotherapy;
KW topoisomerase inhibitor; ss.
XX
OS Homo sapiens.
XX
```

FN W09915648-A1.
XX
PD 01-APR-1999.
XX
PF 17-SEP-1998; 98WO-GB02820.
XX
XX 06-JUN-1998; 98GB-0012140.
PR 23-SEP-1997; 97GB-0020107.
PR 17-OCT-1997; 97GB-0022012.
XX
PA (ISIS-) ISIS PHARM INC.
PA (ZENE) ZENECA LTD.
XX
PI Dean NM, Koropatnick DJ, Vincent MD;
XX
DR WPI; 1999-254708/21.
XX
PT New antisense oligonucleotides directed against thymidylate synthase
XX
PS Disclosure; Fig 7; 53pp; English.
XX
CC This invention describes novel antisense oligonucleotides targeted to
CC sequences in the 3' end of thymidylate synthase (TS) mRNA. Such
CC oligonucleotides are cytosatic on their own when administered to
CC human tumour cell lines, and also enhance the toxicity of anticancer
CC drugs such as Tomudex administered to those cells. In addition, antisense
CC oligonucleotides targeted to 5' sequences induce TS gene transcription.
CC The antisense oligonucleotides are used in a method for the
CC treatment of cancer (or a method for providing antiproliferative effect)
CC The antisense oligonucleotides are also used in the production of a
CC medicant for the treatment of cancer, either separately or in conjunction
CC with a therapeutic agent such as thymidylate synthase (TS) inhibitors
CC (e.g. Tomudex, Zeneca development compound ZD9331 etc.); cytostatic
CC agents (e.g. antioestrogens, (anti)progestogens, antiandrogens,
CC testosterone inhibitors, anti-invasion agents, growth factor inhibitors,
CC etc.); antiproliferative/antineoplastic agents (e.g. antimetabolites,
CC antitumor antibiotics, alkylating agents, antimitotic agents,
CC topoisomerase inhibitors, etc.), or radiotherapy.
XX
SQ Sequence 566 BP; 156 A; 101 C; 135 G; 174 T; 0 other;

Query Match 5.1%; Score 60; DB 20; Length 566;
Best Local Similarity 100.0%; Pred. No. 6.6e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 900 GCCTCCGTCCTCCCGCCGCGCCATGCTGTGCGCGCTCGAGCTCCCGCCGCC 959
|||||
Db 1 GCCTCCGTCCTCCCGCCGCGCCATGCTGTGCGCGCTCGAGCTCCCGCCGCC 60
|||||

Search completed: November 29, 2002, 00:35:19
Job time : 259.395 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

M nucleic - nucleic search, using sw model

Run on: November 28, 2002, 22:48:55 ; Search time 3021.34 Seconds
(without alignments)
11433.695 Million cell updates/sec

Title: US-09-963-333-7

Perfect score: 1187

Sequence: 1 gatcgccactgcactoca.....gagggaggcggtgsgga 1187

Scoring table:

OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 15

Total number of hits satisfying chosen parameters: 471885

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	658	55.4	1186	9	HUMTS	D00517 Homo sapien
2	658	55.4	18596	6	AR144965	AR144965 Sequence
3	658	55.4	18596	6	AX050451	AX050451 Sequence
4	658	55.4	18596	6	AX330682	AX330682 Sequence
5	658	55.4	18596	6	AX330906	AX330906 Sequence
6	658	55.4	18596	6	AX335755	AX335755 Sequence
7	658	55.4	18596	6	AX397714	AX397714 Sequence
8	658	55.4	18596	6	AX408943	AX408943 Sequence
9	658	55.4	18596	9	HUMTS1	D00596 Homo sapien
10	545	45.9	152711	9	AP001178	AP001178 Homo sapi
c 11	332	28.0	165742	2	AC020697	AC020697 Homo sapi
12	297	25.0	1536	6	AR103889	AR103889 Sequence
13	297	25.0	1536	6	AX397715	AX397715 Sequence
14	297	25.0	1536	6	AX482020	AX482020 Sequence
15	297	25.0	1536	9	HSTSYN1	X02308 Human mRNA
16	274	23.1	1567	9	AB062290	AB062290 Homo sapi
17	274	23.1	1569	9	BC013919	BC013919 Homo sapi
18	273	23.0	1533	9	BC002567	BC002567 Homo sapi
19	262	22.1	3298	6	AX281791	AX281791 Sequence
20	205	17.3	693	9	AB077207	AB077207 Homo sapi
21	205	17.3	840	9	AB077208	AB077208 Homo sapi
22	205	17.3	942	6	AX069365	AX069365 Sequence
23	205	17.3	942	6	AX397717	AX397717 Sequence
24	170	14.3	249	9	AF279905	AF279905 Homo sapi
c 25	169	14.2	1133	6	AX397748	AX397748 Sequence
c 26	165	13.9	860	9	BC028301	BC028301 Homo sapi
c 27	165	13.9	1131	6	AX397698	AX397698 Sequence
c 28	165	13.9	1161	6	AX397684	AX397684 Sequence
29	142	12.0	165	9	AF279907	AF279907 Homo sapi
30	142	12.0	193	9	AF279906	AF279906 Homo sapi
31	110	9.3	168	9	AF127520	AF127520 Homo sapi
32	108	9.1	535	6	AX341901	AX341901 Sequence
33	82	6.9	140	9	AF127519	AF127519 Homo sapi
c 34	69	5.8	142629	30	AC027512	AC027512 Homo sapi
c 35	69	5.8	159349	9	AC097633	AC097633 Homo sapi
c 36	69	5.8	160693	9	AC124043	AC124043 Homo sapi
c 37	68	5.7	39958	9	AC011542	AC011542 Homo sapi
c 38	68	5.7	104600	9	AP001166	AP001166 Homo sapi
39	68	5.7	107480	9	AL596094	AL596094 Human DNA
40	68	5.7	136283	9	AP005202	AP005202 Homo sapi
c 41	68	5.7	208145	2	AC053504	AC053504 Homo sapi
42	67	5.6	335	9	AF134215	AF134215 Homo sapi
43	64	5.4	176521	9	AC092724	AC092724 Homo sapi
c 44	63	5.3	1938	9	AB060870	AB060870 Macaca fa
c 45	63	5.3	119330	9	AL590076	AL590076 Human DNA

ALIGNMENTS

RESULT 1

HUMTS

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

HUMTS 1186 bp DNA linear PRI 06-NOV-2001
Homo sapiens gene for thymidylate synthase, exon 1, partial cds.

D00517

D00517.1 GI:220133

Homo sapiens DNA, clone:lambda HTS-1.

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Rutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1186)

Takeishi,K., Kaneda,S., Ayusawa,D., Shimizu,K., Gotch,O. and

Seno,T.

Human thymidylate synthase gene: isolation of phage clones which

cover a functionally active gene and structural analysis of the region upstream from the translation initiation codon
J. Biochem. 106 (4), 575-583 (1989)

source 1. .18596
/organism="unknown"
USE COUNT 4521 a 3991 c 4479 g 5605 t
ORIGIN

Query Match 55.4%; Score 658; DB 6; Length 18596;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

1 GATCGCGCACTGCTCCAGCGCTGGGTGAGAGAGCGAGACTCTGTCTCAAAAAA 60
78 GATCGCGCACTGCTCCAGCGCTGGGTGAGAGAGCGAGACTCTGTCTCAAAAAA 137
61 AAAAGAGCGCGCGGCTCAAAACAAAAACCTCGAAAAAGCCCTGGCGGTCTTTT 120
138 AAAAGAGCGCGCGGCTCAAAACAAAAACCTCGAAAAAGCCCTGGCGGTCTTT 197
121 TTTTCTTTTCTTTTCTTTTGGGACAGTCTTCTCTCTGCGCCAGGCTGAGTACAT 180
198 TTTTCTTTTCTTTTCTTTTGGGACAGTCTTCTCTCTGCGCCAGGCTGAGTACAT 257
181 CGGATCTTGGCTCACTGCAACTCTGCTTCCAGGTTCAAGCAATCTTCTGCTCAG 317
241 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTATTTTGTAGTACAGCGGGTTTCA 300
318 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTATTTTGTAGTACAGCGGGTTTCA 377
301 CATGTTGTCCAGGCTGCTCTGCACTCTGCACTCTGCACTCTGCACTCTGCACT 360
378 CATGTTGTCCAGGCTGCTCTGCACTCTGCACTCTGCACTCTGCACTCTGCACT 437
361 CAAGTACTAGGATTACAGCGGTGAGCACCAGCTCCAGCGCTCCAGCGCTCCAG 420
438 CAAGTACTAGGATTACAGCGGTGAGCACCAGCTCCAGCGCTCCAGCGCTCCAG 497
421 AGTAAAGAGCTGATTTATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
498 AGTAAAGAGCTGATTTATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556
481 AAATGCAATCCTTTATGTTAGTAAACAGATCTCAAAACAGCAGTCTTTGTCAG 540
557 AAATGCAATCCTTTATGTTAGTAAACAGATCTCAAAACAGCAGTCTTTGTCAG 616
541 ACCGAGGAAACCTGGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
617 ACCGAGGAAACCTGGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
601 TTCCCAAAGGCGCAGTCTTCCGAGCAGCAGCTGCTGCTGCTGCTGCTGCTGCT 660
677 TTCCCAAAGGCGCAGTCTTCCGAGCAGCAGCTGCTGCTGCTGCTGCTGCTGCT 736
661 AAGACTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
737 AAGACTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
721 CCCCTGGCGCAGCTCTTCTAGAGCGGGGCGCGCGAGCCCGCGAGAGAGAGGG 780
797 CCCCTGGCGCAGCTCTTCTAGAGCGGGGCGCGCGAGCCCGCGAGAGAGAGGG 856
781 GAGCGCGGAGCGCGCGGAAAGAGCGCGGAGAGGGGCTCTGCTGCTGCTGCTGCT 840
857 GAGCGCGGAGCGCGCGGAAAGAGCGCGGAGAGGGGCTCTGCTGCTGCTGCTGCT 916
841 CTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
917 CTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
901 CTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
977 CTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036

961	TTGCCCCCGCCGACACAGAGCGGGAGCGCGAGCGCGGCTCCGCGCAGCGGGAGCTCCAG	1020	421	AGTGAAGAAAGCTGCATTATACCACTTCCTTCGCTTCGCTTCAGTTCAGAAACGAAATGG	480
1037	TTGCCCCCGCCGACACAGAGCGGGAGCGCGAGCGCGCTCCGCGCAGCGGGAGCTCCAG	1096	498	AGTGAAGAAAGCTGCATTATACCACTTCCTTCGCTTCGCTTCAGTTCAGAAACGAAATGG	556
1021	TACCTGGGGAGAGTCCACACATCTCCGCTGGCGGTGAGGAAGAGCAGCCGACCGGC	1080	481	AAATGCAAATCTCTATTAGTTGTAGGAACAGATCTCAACAGCAGTTCCTTCGCAAG	540
1097	TACCTGGGGAGAGTCCACACATCTCCGCTGGCGGTGAGGAAGAGCAGCCGACCGGC	1156	557	AAATGCAAATCTCTATTAGTTGTAGGAACAGATCTCAACAGCAGTTCCTTCGCAAG	616
1081	ACCGGACCTCTGCGGTATTCGGCATGACGCGCGCTTACAGCTGAGAGGTGACGCGCG	1140	541	ACCGGACAAAGCTGGGAACCTGTCTGCTGTAGGAAGAGCGCGGTTCACAGACGG	600
1157	ACCGGACCTCTGCGGTATTCGGCATGACGCGCGCTTACAGCTGAGAGGTGACGCGCG	1216	617	ACCGGACAAAGCTGGGAACCTGTCTGCTGTAGGAAGAGCGCGGTTCACAGACGG	676
1141	GGCCCTTGGCGGACCGGCTGGGAGGAGGAGGCGCGCTGGGGA	1187	601	TTCCCAAGGCGCGAGTCTCTCCGCTTCGCTTCAGGCTTCAGGTTCCCGGCTTCCT	660
1217	GGCCCTTGGCGGACCGGCTGGGAGGAGGAGGCGCGCTGGGGA	1263	677	TTCCCAAGGCGCGAGTCTCTCCGCTTCGCTTCAGGCTTCAGGTTCCCGGCTTCCT	736
RESULT 5			661	AAGACTCTCAGCTGTGGCTGGGCTCCGCTTCGTGTGCGCACACCCGTGCTTCCTCGCTTC	720
LOCUS	AX330906	18596 bp	737	AAGACTCTCAGCTGTGGCTGGGCTCCGCTTCGTGTGCGCACACCCGTGCTTCCTCGCTTC	796
DEFINITION	Sequence 1415 from Patent WO0194629.		721	CCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGACCCCGCGCAGCAGAGAGGCG	780
ACCESSION	AX330906		797	CCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGACCCCGCGCAGCAGAGAGGCG	856
VERSION	AX330906.1	GI:18103885	781	GAGCGCGGACGCGCCGCGGAAAGCGCGCGGAGGGTCTTGCAACCGCGCACTTGG	840
KEYWORDS	human.		857	GAGCGCGGACGCGCCGCGGAAAGCGCGCGGAGGGTCTTGCAACCGCGCACTTGG	916
SOURCE	human.		841	CCTGCTCCGCTCCGCGCGCGCACCTTGGCTTCGCTTCGCTTCGCGCGCCTTCGCTG	900
ORGANISM	Homo sapiens		917	CCTGCTCCGCTCCGCGCGCGCACCTTGGCTTCGCTTCGCTTCGCGCGCCTTCGCTG	976
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		901	CCTGCTCCGCTCCGCGCGCGCACCTTGGCTTCGCTTCGCTTCGCGCGCCTTCGCTG	960
AUTHORS	Young, P.B., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horrigan, S., Soppet, D.R. and Weaver, Z.		977	CCTGCTCCGCTCCGCGCGCGCACCTTGGCTTCGCTTCGCTTCGCGCGCCTTCGCTG	1036
TITLE	Cancer gene determination and therapeutic screening using signature gene sets		961	TTGCCCCCGCCGACACAGAGCGGGAGCGCGCTCCGCGCAGCGGGAGCTGCGAG	1020
JOURNAL	Patent: WO 0194629-A 1415 13-DEC-2001;		1037	TTGCCCCCGCCGACACAGAGCGGGAGCGCGCTCCGCGCAGCGGGAGCTGCGAG	1096
FEATURES	Avalon Pharmaceuticals (US)		1021	TACCTGGGGAGAGTCCACACATCTCTCCGCTGGCGGTGAGGAAGAGCAGCGCGGC	1080
source	Location/Qualifiers		1097	TACCTGGGGAGAGTCCACACATCTCTCCGCTGGCGGTGAGGAAGAGCAGCGCGGC	1156
BASE COUNT	4521 a 3991 c 4479 g 5605 t		1081	ACCGGACCTCTGCGGTATTCGGCATGACGCGCGCTTACAGCTGAGAGGTGACGCGCG	1140
ORIGIN			1157	ACCGGACCTCTGCGGTATTCGGCATGACGCGCGCTTACAGCTGAGAGGTGACGCGCG	1216
Query Match	55.4%; Score 658; DB 6; Length 18596;		1141	GGCCCTTGGCGGACCGGCTGGGAGGAGGAGGCGCGCTGGGGA	1187
Best Local Similarity	99.2%; Pred. No. 0;		1217	GGCCCTTGGCGGACCGGCTGGGAGGAGGAGGCGCGCTGGGGA	1263
Matches 1178; Conservative	0; Mismatches 8; Indels 1; Gaps 1;		RESULT 6		
1	GATCGGCCACTGCACTCCAGCTGGGTGAGAGAGGAGACTCTGTCTCAAAAAA	60	AX335755	18596 bp	DNA
78	GATCGGCCACTGCACTCCAGCTGGGTGAGAGAGGAGACTCTGTCTCAAAAAA	137	LOCUS	AX335755	Sequence 6264 from Patent WO0194629.
61	AAAAAGACCGCAGGCTCAACAAAAACCTCGAAAAAGCCTGGCGGTCTTTTTTTT	120	DEFINITION	AX335755	
138	AAAAAGACCGCAGGCTCAACAAAAACCTCGAAAAAGCCTGGCGGTCTTTTTTTT	197	ACCESSION	AX335755.1	GI:18126474
121	TTTTTTTTTTTTTTTTTTGGACAGTCTGCTCTGTCGCCAGGCTGGAGTACAAAGGT	180	VERSION	human.	
198	TTTTTTTTTTTTTTTTTTGGACAGTCTGCTCTGTCGCCAGGCTGGAGTACAAAGGT	257	KEYWORDS	human.	
181	CGGATCTTGCTCACTCGCAACTCTGCTCCAGGTTCAAGCAATTCCTGCTCCAGCC	240	SOURCE	Homo sapiens	
258	CGGATCTTGCTCACTCGCAACTCTGCTCCAGGTTCAAGCAATTCCTGCTCCAGCC	317	ORGANISM	Homo sapiens	
241	TCCCAAGTAGCCACACGCGCCAGCTAATTTTTGTANTTTTTAGTAGAGCGGGGTTTCAC	300	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
318	TCCCAAGTAGCCACACGCGCCAGCTAATTTTTGTANTTTTTAGTAGAGCGGGGTTTCAC	377	AUTHORS	Young, P.B., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horrigan, S., Soppet, D.R. and Weaver, Z.	
301	CATGTTGTCAGGCTGGTCTGTAACCTCTGACCTCAGGTCATCCACCGCTTCGCGCC	360	TITLE	Cancer gene determination and therapeutic screening using signature gene sets	
378	CATGTTGTCAGGCTGGTCTGTAACCTCTGACCTCAGGTCATCCACCGCTTCGCGCC	437	JOURNAL	Patent: WO 0194629-A 6264 13-DEC-2001;	
361	CAAGTACTAGGATTCAGGCTGAGCAGCGCTTCAGCGCTTCAGCGCTTCGCGCTTAAATCA	420	FEATURES	Avalon Pharmaceuticals (US)	
438	CAAGTACTAGGATTCAGGCTGAGCAGCGCTTCAGCGCTTCAGCGCTTCGCGCTTAAATCA	497		Location/Qualifiers	

438	ib	CAAAGTACTAGAGTATTACAGCGTGAGCCACACGCGTCCAGCGCCCTTGGCGGTTTTTAATCA	497
440			
442	iy	AGTAGAAAAGCTGCATTATACCACTTGTCTTCGTTGCTTCAGTGTAGAAACGAGAAATGG	480
444			
446	ib	AGTAGAAAAGCTGCATTATACCACTTGTCTTCGGTTGC-TTCAGTGTAGAACGAGAAATGG	556
448			
450	iy	AAATGCAAAATCNCNTTATTAGTGTGTAGSAAACAGATCTCAAACAGCAGTTTTTTGTNGACAAG	540
452			
454	ib	AAATGCAAAATCCCTTATTAGTGTGTAGSAAAACAGATCTCAAACAGCAGTTTGTGTGACAG	616
456			
458	iy	ACCGAGAGAAAACGTGGGAACTGTGTCTGTGTGTAGAGAGGCGCGTGTGACAGACGG	600
460			
462	ib	ACCGAGAGAAAACGTGGGAACTGTGTCTGTGTGTAGAGAGGCGCGTGTGACAGACGG	676
464			
466	iy	TTCCCAAAGGGGCGAGTGCCTTCCCGGCCACCGCACTCTGCACAGTTCCCGGGTTCCT	660
468			
470	ib	TTCCCAAAGGGGCGAGTGCCTTCCCAAGCCACCGCACTCTGCACAGTTCCCGGGTTCCT	736
472			
474	iy	AAGACTCTCAGCTGTGCGCCCTGGGCTTCTGTGTGCAACCCCTGTGCTCTTCGCTTTC	720
476			
478	ib	AAGACTCTCAGCTGTGCGCCCTGGGCTTCTGTGTGCAACCCCTGTGCTCTTCGCTTTC	796
480			
482	iy	CCCTGTGCGCACGCTCTTAGAGCGGGGCGCGCGCACCCCGCCGAGCAAGAGAGCG	780
484			
486	ib	CCCTGTGCGCACGCTCTTAGAGCGGGGCGCGCGCACCCCGCCGAGCAAGAGAGCG	856
488			
490	iy	GAGCGCGGACGCGCCGCGGAAAGCGCGCGSAGGGGTCTGCCACCGCGCCACTTGG	840
492			
494	ib	GAGCGCGGACGCGCCGCGGAAAGCGCGSAGGGGTCTGCCACCGCGCCACTTGG	916
496			
498	iy	CTTGCTTCGTCTCGCGCGCGCACCTTGGCGTTCCTCGTCTCGCGCGCGCACCTTGGCGTG	900
500			
502	ib	CTTGCTTCGTCTCGCGCGCGCACCTTGGCGTTCCTCGTCTCGCGCGCGCACCTTGGCGTG	976
504			
506	iy	CTTCTGTCCCCCGCGCGCGCGCATGTGCTGTGCGCGGTCTCGAGCTGCGCGCGCGCGCC	960
508			
510	ib	CTTCTGTCCCCCGCGCGCGCGCATGTGCTGTGCGCGGTCTCGAGCTGCGCGCGCGCGCC	1036
512			
514	iy	TTTGCCCCCGCGCGCACGAGAGCGGACCGCGAGCGCGTCCGCGCGACGCGGAGCTGCGAG	1020
516			
518	ib	TTTGCCCCCGCGCGCACGAGAGCGGACCGCGAGCGCGTCCGCGCGACGCGGAGCTGCGAG	1096
520			
522	iy	TACTGTGGGCGAGATCCAAACACATCTCTCGCTGTGGGCTCAGGAAGSACGACCGCACGGCG	1080
524			
526	ib	TACTGTGGGCGAGATCCAAACACATCTCTCGCTGTGGGCTCAGGAAGSACGACCGCACGGCG	1156
528			
530	iy	ACCGSACCCCTGTTCGCTGTATTCGCGCATGTGCGCATGTAGCGCGCTACAGCTGTAGAGGTGTGCGCGCG	1140
532			
534	ib	ACCGSACCCCTGTTCGCTGTATTCGCGCATGTGCGCATGTAGCGCGCTACAGCTGTAGAGGTGTGCGCGCG	1216
536			
538	iy	GGGCCCCCTGCGGGACGGGTGTGCGGGAAAGAGAGGCGCGGTGTGGGGA	1187
540			
542	ib	GGGCCCCCTGCGGGACGGGTGTGCGGGAAAGAGAGGCGCGGTGTGGGGA	1263

RESULT 8
 LOCUS AX408943
 DEFINITION Sequence 1590 bp DNA linear PAT 14-JUN-2002
 ACCESSION AX408943
 VERSION AX408943.1
 KEYWORDS GI:21441648
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 REFERENCE
 AUTHORS Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
 TITLE Gene expression profiles in liver cancer
 JOURNAL Patent: WO 0229103-A 1590 11-APR-2002;
 GENE LOGIC INC (US)
 FEATURES
 source Location/Qualifiers
 1..18596

/organism="Homo sapiens"	
/db xref="taxon:9606"	
/note="EMBL/GenBank Accession No. D00596"	
BASE COUNT	4521 a 3991 c 4479 g 5605 t
ORIGIN	
Query Match	55.4%; Score 658; DB 6; Length 18596;
Best Local Similarity	99.2%; Pred. No. 0;
Matches 1178; Conservative	0; Mismatches 8; Indels 1; Gaps
Qy	1 GATCGCGCACTGCACTCCAGCCTGGGTGAGAGCGAGACTCTCTCTCAAAAAA 60
Db	78 GATCGCGCACTGCACTCCAGCCTGGGTGAGAGCGAGACTCTCTCTCAAAAAA 137
Qy	61 AAAAAGACCGCAGCGGCTCAAAACAAAACCTCGSAAAAGCCCTCGGCGTCTTTTTTT 120
Db	138 AAAAAGACCGCAGCGGCTCAAAACAAAACCTCGSAAAAGCCCTCGGCGTCTTTTTTT 197
Qy	121 TTTTTTTTTTTTTTTTTTGGGACAGTCTTGCTCTGTGCGCCAGCTGGAGTCAATGGT 180
Db	198 TTTTTTTTTTTTTTTTTTGGGACAGTCTTGCTCTGTGCGCCAGCTGGAGTCAATGGT 257
Qy	181 CGGATCTTGCGTCACTGCAACCTCTGCTCCACGGTTCAAGCAATTTCTGTGCTCAGCC 240
Db	258 CGGATCTTGCGTCACTGCAACCTCTGCTCCACGGTTCAAGCAATTTCTGTGCTCAGCC 317
Qy	241 TCCCAAGTAGCACACACGCCACCTAATTTTGTATTTTACTAGAGACGGGGTTTCAC 300
Db	318 TCCCAAGTAGCACACACGCCACCTAATTTTGTACTTTTAGTAGACGGGGTTTCAC 377
Qy	301 CATGTTCTCAGCGTGCTGTAAGCTCTGTGACCTCAGGTGATCCACCGCTCGGCCGCC 360
Db	378 CATGTTCTCAGCGTGCTGTAAGCTCTGTGACCTCAGGTGATCCACCGCTCGGCCGCC 437
Qy	361 CAAGAATCTAGATTTACAGCGCTGAGCCACCGCTCCAGCGCTCTGGCGGTTTTTAATCA 420
Db	438 CAAGAATCTAGATTTACAGCGCTGAGCCACCGCTCCAGCGCTCTGGCGGTTTTTAATCA 497
Qy	421 AGTAGAAAAAGCTGCATATACCACTTGCTTCNGTTGNTTTCAGTCAGAACGAAATGG 480
Db	498 AGTAGAAAAAGCTGCATATACCACTTGCTTCGGTTGC-TTCAGTCAGAACGAAATGG 556
Qy	481 AAATGCAAAATCNCCTTATAGTTTAGGAAAAACAGATCTCAAAACAGCAGTTTGTGACAG 540
Db	557 AAATGCAAAATCNCCTTATAGTTTAGGAAAAACAGATCTCAAAACAGCAGTTTGTGACAG 616
Qy	541 ACCCGAGGAAACGTGGGAACTGTGCTGGCTTAGAGAGGCGCGGTGACACAGCG 600
Db	617 ACCCGAGGAAACGTGGGAACTGTGCTGGCTTAGAGAGGCGCGGTGACACAGCG 676
Qy	601 TTCCCAAGGGCGAGTCTTTCCNGCCACACGACCTGNTCCAGTTCCCGGTTTCT 660
Db	677 TTCCCAAGGGCGAGTCTTTCCACAGCACCGACCTGCATCCAGTTCCCGGTTTCT 736
Qy	661 AAGACTCTCAGCTGTGGCCCTGGGCTCCGTTCTGTGCCACACCCGTGGCTCTCGGTTTC 720
Db	737 AAGACTCTCAGCTGTGGCCCTGGGCTCCGTTCTGTGCCACACCCGTGGCTCTCGGTTTC 796
Qy	721 CCCCTGCGCACGCTCTCTAGACGGGGGCGCGCGACCCCGCGACGAGGAAGGCG 780
Db	797 CCCCTGCGCACGCTCTCTAGACGGGGGCGCGCGACCCCGCGACGAGGAAGGCG 856
Qy	781 GAGCGGGGACGCGCGGGAAAGCGCGCGGAAGGGGTCGTGCCACCGCGCCACTTGG 840
Db	857 GAGCGGGGACGCGCGGGGAAAGSGCGCGGAAGGGGTCGTGCCACCGCGCCACTTGG 916
Qy	841 CTTGCTCTGTCGCGCGCGCACCTTGGGCTGCTCGTCCCGCGCGGCCACTTTCGCTG 900
Db	917 CTTGCTCTGTCGCGCGCGCACCTTGGGCTGCTCGTCCCGCGCGGCCACTTTCGCTG 976
Qy	901 CTTGCTCTCCCGCGCGCGCACCTGCTGTGGCGGGCTCGGAGTTCGCGCGCGGCC 960
Db	977 CTTGCTCTCCCGCGCGCGCGCACCTGCTGTGGCGGGCTCGGAGTTCGCGCGCGGCC 1036

QY 961 TTGCCCCCGCGGACAGGAGCGGAGCCGAGCGCGTCCGCGCACGGGAGCTGCAG 1020
 Db 1037 TTGCCCCCGCGGACAGGAGCGGAGCGCGTCCGCGCACGGGAGCTGCAG 1096
 QY 1021 TACTCTGGGCGAGATCCACACATCTCTCGCTGCGGCGTACAGGAGCGACCGCGCGC 1080
 Db 1097 TACTCTGGGCGAGATCCACACATCTCTCGCTGCGGCGTACAGGAGCGACCGCGCGC 1156
 QY 1081 ACCGCAACCTCTGCGTATTCGGCATGTCAGGCGCGCTACAGCTTGAGAGTGCACGCCGCG 1140
 Db 1157 ACCGCAACCTCTGCGTATTCGGCATGTCAGGCGCGCTACAGCTTGAGAGTGCACGCCGCG 1216
 QY 1141 GCGCCCTCGGAGCGGTCGCGGAGGAGGCGCGCGTGGGGA 1187
 Db 1217 GCGCCCTCGGAGCGGTCGCGGAGGAGGCGCGCGTGGGGA 1263

RESULT 9

HUMTS1 18596 bp DNA linear PRI 14-APR-2000
 LOCUS Homo sapiens gene for thymidylate synthase, exons 1, 2, 3, 4, 5, 6,
 DEFINITION 7, complete cds.

ACCESSION D00596
 VERSION D00596.1 GI:220135
 KEYWORDS thymidylate synthase.
 SOURCE Homo sapiens DNA, clone:lambdaHTS-1 and lambdaHTS-3.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 18596)
 AUTHORS Kaneda, S., Nalbantoglu, J., Takeishi, K., Shimizu, K., Gotoh, O.,
 Sero, T. and Ayusawa, D.

TITLE Structural and functional analysis of the human thymidylate
 synthase gene

JOURNAL J. Biol. Chem. 265 (33), 20277-20284 (1990)

MEDLINE 91056070
 COMMENT These data kindly submitted in computer readable form by: Sumiko
 Kaneda

National Institute of Genetics

1111 Yata

Mishima 411

Japan

Phone: +81-559-72-2732

Fax: +81-559-71-3651.

Location/Qualifiers

1. .18596

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="18"

/map="18p11.32"

/clones="lambdaHTS-1 and lambdaHTS-3"

1. .148

/note="Alu sequence"

202. .477

/note="Alu sequence"

prim_transcript 822. .16246

/note="thymidylate synthase mRNA and introns"

prim_transcript 824. .16246

/note="thymidylate synthase mRNA and introns"

prim_transcript 828. .16246

/note="thymidylate synthase mRNA and introns"

prim_transcript 841. .16246

/note="thymidylate synthase mRNA and introns"

repeat_unit 862. .889

/note="inverted repeat"

904. .993

/note="triple tandem repeats"

join(1001..1205,2895..2968,5396..5570,11843..11944,

13449..13624,14133..14204,15613..15750)

/EC number="2.1.1.45"

/codon_start=1

/product="thymidylate synthase"

/protein_id="BAA00472.1"
 /db_xref="GI:220136"
 /translation="MFVAGSELPRRLPPAAQERDAEPHPHGLQYLGQIHLRGG
 VRKDDRTGTLSVFGQARYSLRDEFPILTKRVFKWGVFVLELLWFKGSTNAKLS
 SKVKIDANGSRDPLDILGFSREBGLGVYGFQWRFVGAERDMESDYGQGVQ
 LORVIDITKPNDDRRIIMCAWNPRLPLMALPECHALCOFYVNSLSCLYQSRSD
 MGLVPPNTIASALLTYMLAHITGLKPGDFHTLGDAAHYIYNHIEPLIKIQREPRPF
 PKGLIRKVEKIDDFKABDFOLEGYNPHPTIKMEMAV"
 <1001..1205
 exon /number=1
 1206..2894
 intron /number=1
 1660..1665
 protein_bind /bound moiety="Sp1"
 1890..1895
 protein_bind /bound moiety="Sp1"
 1895..1900
 protein_bind /bound moiety="Sp1"
 2895..2968
 exon /number=2
 2969..5395
 intron /number=2
 3049..3347
 repeat_unit /note="Alu sequence"
 4953..5245
 repeat_unit /note="Alu sequence"
 5396..5570
 exon /number=3
 5571..11842
 intron /number=3
 5768..9405
 repeat_unit /note="L1 sequence"
 10125..10389
 repeat_unit /note="reiterated motif of GGA/TCAT"
 11843..11944
 exon /number=4
 11945..13448
 intron /number=4
 12135..12397
 repeat_unit /note="Alu sequence"
 12567..12743
 repeat_unit /note="Alu sequence"
 12807..13084
 repeat_unit /note="Alu sequence"
 13449..13624
 exon /number=5
 13625..14132
 intron /number=5
 13875..14010
 repeat_unit /note="Alu sequence"
 14133..14204
 exon /number=6
 14205..15612
 intron /number=6
 14531..14832
 repeat_unit /note="Alu sequence"
 15613..15750
 exon /number=7
 15934..15939
 polyA_signal
 16228..16233
 polyA_signal
 17262..17553
 repeat_unit /note="Alu sequence"
 BASE COUNT 4521 a 3991 c 4473 g 5605 t
 ORIGIN

Query Match 55.4%; Score 658; DB 9; Length 18596;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1 GATCGCGCCACTGCACCTCCAGCTGGGTGGAGAGCGAGACTCTCTCTCAAAAAA 60

|||||

Db 78 GATCGCGCCACTGCACCTCCAGCTGGGTGGAGAGCGAGACTCTCTCTCAAAAAA 137

|||||

Db 1217 GGCCCCGCGGGACGGGTGGCGGAAGAGGAGCGCGCTGTGGGA 1263
|||||

RESULT 10
AP001178 152711 bp DNA linear PRI 11-JUL-2002
LOCUS Homo sapiens genomic DNA, chromosome 18 clone:RP11-806L2, complete sequence.

DEFINITION AP001178 GI:21728142
HTG.
AP001178.4 Homo sapiens DNA, clone:RP11-806L2.

ACCESSION AP001178
VERSION AP001178.4
KEYWORDS HTG.

SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2000)

REFERENCE 2 (bases 1 to 152711)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE Direct Submission
JOURNAL Submitted (14-FEB-2000) Mashira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://bgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)

COMMENT On Jul 10, 2002 this sequence version replaced gi:9229948.

FEATURES
Location/Qualifiers
1..152711
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="18"
 /map="18p"
 /clone="RP11-806L2"

BASE COUNT 42612 a 32833 c 33397 g 43869 t

ORIGIN
Query Match 45.9%; Score 545; DB 9; Length 152711;
Best Local Similarity 99.2%; Pred. No. 8.8e-298;
Matches 1065; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 114 TTTTNTTTTTTTTTTTTTTTTTTTTTTTTGGAAGATCTTCGTGCTGCCAGGCTGGAGTA 173
|||
Db 84619 TTTTNTTTTTTTTTTTTTTTTTTTTTTTTGGAAGATCTTCGTGCTGCCAGGCTGGAGTA 84678
|||

QY 174 CAATGGTCCGATCTTGGCTCACTGCACACTCTGCTCCCAGGTTCAAGCAATTCTTCGC 233
|||
Db 84679 CAATGGTCCGATCTTGGCTCACTGCACAACTCTGCTCCCAGGTTCAAGCAATTCTTCGC 84738
|||

QY 234 CTCAGCCTCCCAAGTAGCCACACCGCCAGCTAATTTTGTANTTTTAGTAGACGGGG 293
|||
Db 84739 CTCAGCCTCCCAAGTAGCCACCACCGCCAGCTAATTTTGTACTTTTAGTAGACGGGG 84798
|||

QY 294 GTTTCACCATGTTGTCAGGCTGGCTNGAACTCTGACCTCAGGTGATCCACCGGCTC 353
|||
Db 84799 GTTTCACCATGTTGTCAGGCTGGCTNGAACTCTGACCTCAGGTGATCCACCGGCTC 84858
|||

QY 354 GGCCCCCAAAGTACTAGGATTCAGGCGTGAGCCACCGCGTCCAGCGCCTGCGGTTT 413
|||
Db 84859 GGCCCCCAAAGTACTAGGATTCAGGCGTGAGCCACCGCGTCCAGCGCCTGCGGTTT 84918
|||

QY 414 TTAAATCAAGTAGAAAAGCTGCATTATACACTTGCCTTCGTTTCGTTTCAGTGAGAACGAA 473
|||
Db 84919 TTAAATCAAGTAGAAAAGCTGCATTATACACTTGCCTTCGTTTCGTTTCAGTGAGAACGAA 84977
|||

QY 474 GAATCGGAATGCAATCNCTTATTAGTTGTAGGAAACAGATCTCTAAACAGCAGTTTGT 533
|||
Db 84978 GAATCGGAATGCAATCCCTTATTAGTTGTAGGAAACAGATCTCTAAACAGCAGTTTGT 85037
|||

[illegible]

Db 14 CCACCTTGGCTTGCCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 73
QY 893 TTGCGCTGCTTCCGTCCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 952
Db 74 TTGCGCTGCTTCCGTCCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 133
QY 953 GCGCGCCCTTGGCTTCCGTCCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 1012
Db 134 GCGCGCCCTTGGCTTCCGTCCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 193
QY 1013 AGCTGAGTACCTGGGCGAGATCCACACATCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 1072
Db 194 AGCTGAGTACCTGGGCGAGATCCACACATCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 253
QY 1073 GCACGGGACCGGACCGCTGCTGCTTTCGGCATGCGGCGGCTTGGCTTGCCTCCGTCGCCGCGCCAC 1129
Db 254 GCACGGGACCGGACCGCTGCTGCTTTCGGCATGCGGCGGCTTGGCTTGCCTCCGTCGCCGCGCCAC 310

RESULT 13
AX397715 1536 bp DNA linear PAT 18-MAY-2002
LOCUS
DEFINITION Sequence 141 from Patent WO0210366.
ACCESSION AX397715
VERSION AX397715.1 GI:21068437
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Chen, H., Freimer, N.B. and Novak, T.
TITLE Methods and compositions for diagnosing and treating chromosome-18p related disorders
JOURNAL Patent: WO 0210366-A 141 07-FEB-2002;
Millennium Pharmaceuticals, Inc. (US); The Regents of The University of California (US)
FEATURES
Location/Qualifiers
1..1536
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 390 a 369 c 399 g 378 t
ORIGIN

Query Match 25.0%; Score 297; DB 6; Length 1536;
Best Local Similarity 100.0%; Pred. No. 4.8e-157;
Matches 297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 833 CCACCTTGGCTTGCCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 892
Db 14 CCACCTTGGCTTGCCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 73
QY 893 TTGCGCTGCTTCCGTCCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 952
Db 74 TTGCGCTGCTTCCGTCCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 133
QY 953 GCGCGCCCTTGGCTTCCGTCCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 1012
Db 134 GCGCGCCCTTGGCTTCCGTCCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 193
QY 1013 AGCTGAGTACCTGGGCGAGATCCACACATCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 1072
Db 194 AGCTGAGTACCTGGGCGAGATCCACACATCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 253
QY 1073 GCACGGGACCGGACCGCTGCTGCTTTCGGCATGCGGCGGCTTGGCTTGCCTCCGTCGCCGCGCCAC 1129
Db 254 GCACGGGACCGGACCGCTGCTGCTTTCGGCATGCGGCGGCTTGGCTTGCCTCCGTCGCCGCGCCAC 310

RESULT 14
AX482020 1536 bp DNA linear PAT 16-AUG-2002
LOCUS

/note="assembly_fragment"
34139..47814
/note="assembly_fragment"
47915..62653
/note="assembly_fragment"
clone end:896
vector side:left
62754..89031
/note="assembly_fragment"
89132..116606
/note="assembly_fragment"
clone end:87
vector side:right
116707..165742
/note="assembly_fragment"
47418 a 35451 c 35457 g 46526 t 800 others

BASE COUNT 47418 a 35451 c 35457 g 46526 t 800 others
ORIGIN

Query Match 28.0%; Score 332; DB 2; Length 165742;
Best Local Similarity 100.0%; Pred. No. 6.7e-177;
Matches 332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 856 CCGCGCCACTTGGCTTCCGTCCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 915
Db 18916 CCGCGCCACTTGGCTTCCGTCCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 18857
QY 916 CCGCGCCACTTGGCTTCCGTCCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 975
Db 18856 CCGCGCCACTTGGCTTCCGTCCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 18797
QY 976 CAGAGCGGAGCGCGAGCGCGTCGCGCGACCGGAGCTCGAGTACCTGGGCGAGATC 1035
Db 18736 CAGAGCGGAGCGCGAGCGCGTCGCGCGACCGGAGCTCGAGTACCTGGGCGAGATC 18737
QY 1036 CAACACATCTCCGTCCTCCGTCGCCGCGCCACCGGAGCTCGAGTACCTGGGCGAGATC 1095
Db 18736 CAACACATCTCCGTCCTCCGTCGCCGCGCCACCGGAGCTCGAGTACCTGGGCGAGATC 18677
QY 1096 GTATTTCGGCATGCGGCGCGCTACAGCTGAGAGTACAGCGCGGCGCCCTTGGCGGACG 1155
Db 18676 GTATTTCGGCATGCGGCGCGCTACAGCTGAGAGTACAGCGCGGCGCCCTTGGCGGACG 18617
QY 1156 GTTGGCGGAGGAGGAGGCGCGCTGGGGA 1187
Db 18616 GTTGGCGGAGGAGGAGGCGCGCTGGGGA 18585

RESULT 12
AR103889 1536 bp DNA linear PAT 14-FEB-2001
LOCUS
DEFINITION Sequence 1 from patent US 6087489.
ACCESSION AR103889
VERSION AR103889.1 GI:12815477
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1536)
AUTHORS Dean, N.M.
TITLE Antisense oligonucleotide modulation of human thymidylate synthase expression
JOURNAL Patent: US 6087489-A 1 11-JUL-2000;
FEATURES
Location/Qualifiers
1..1536
/organism="unknown"

BASE COUNT 390 a 369 c 399 g 378 t
ORIGIN

Query Match 25.0%; Score 297; DB 6; Length 1536;
Best Local Similarity 100.0%; Pred. No. 4.8e-157;
Matches 297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 833 CCACCTTGGCTTGCCTCCGTCGCCGCGCCACCTTGGCTTGCCTCCGTCGCCGCGCCAC 892

DEFINITION Sequence 11 from Patent WO02057489.
AX482020
VERSION AX482020.1 GI:22316743
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Danenberg, K.D.
TITLE Method of determining a chemotherapeutic regimen based on ercc1 and
ts expression
JOURNAL Patent: WO 02057489-A 11 25-JUL-2002;
RESPONSE GENETICS INC (US)
FEATURES
Location/Qualifiers
source 1..1536
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 390 a 369 c 399 g 378 t
ORIGIN
Query Match 25.0%; Score 297; DB 6; Length 1536;
Best Local Similarity 100.0%; Pred. No. 4.8e-157;
Matches 297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 833 CCACCTGGCTGCTCCGTCGCGCGCGCCACCTGGCTGCTCCGTCGCTCCGCGCGCCAC 892
Db 14 CCACCTGGCTGCTCCGTCGCGCGCGCCACCTGGCTGCTCCGTCGCTCCGCGCGCCAC 73
QY 893 TTCGCTGCTCCGTCGCTCCGTCGCGCGCGCCACCTGGCTGCTCCGTCGCTCCGCGCGCCAC 952
Db 74 TTCGCTGCTCCGTCGCTCCGTCGCGCGCGCCACCTGGCTGCTCCGTCGCTCCGCGCGCCAC 133
QY 953 GCCGCGCTTCGCTCCGTCGCGCGCGCCACCTGGCTGCTCCGTCGCTCCGCGCGCCAC 1012
Db 134 GCCGCGCTTCGCTCCGTCGCGCGCGCCACCTGGCTGCTCCGTCGCTCCGCGCGCCAC 193
QY 1013 AGCTGCTGCTGCTCCGTCGCTCCGTCGCGCGCGCCACCTGGCTGCTCCGTCGCTCCGCGCGCCAC 1072
Db 194 AGCTGCTGCTGCTCCGTCGCTCCGTCGCGCGCGCCACCTGGCTGCTCCGTCGCTCCGCGCGCCAC 253
QY 1073 GCACGGGCACCGGACCCCTGCTGGTATTCGGCATGCGAGCGCGCTACAGCTGAGAG 1129
Db 254 GCACGGGCACCGGACCCCTGCTGGTATTCGGCATGCGAGCGCGCTACAGCTGAGAG 310
RESULT 15
HSTS1
LOCUS HSTS1 1536 bp mRNA linear PRI 12-SEP-1993
DEFINITION Human mRNA for thymidylate synthase (EC 2.1.1.45).
ACCESSION X02308
VERSION X02308.1 GI:37478
KEYWORDS inverted repeat; synthetase; tandem repeat.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1536)
AUTHORS Takeishi, K., Kaneda, S., Ayusawa, D., Shimizu, K., Gotoh, O. and
Seno, T.
TITLE Nucleotide sequence of a functional cDNA for human thymidylate
synthase
JOURNAL Nucleic Acids Res. 13 (6), 2035-2043 (1985)
MEDLINE 85215597
PUBMED 2987839
COMMENT Data kindly reviewed (22-OCT-1985) by Seno T.
FEATURES
Location/Qualifiers
source 1..1536
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature 14..103
/notes="triple tandemly repeated elements"
repeat_region 14..16

misc_feature 35..103
/note="direct repeat 1"
misc_feature 35..69
/note="pot. stem-loop structure"
repeat_unit 35..41
/note="pot. stem-loop structure"
repeat_region 42..44
/note="inverted repeat A"
misc_feature 63..103
/note="direct repeat 1"
repeat_unit 63..69
/note="inverted repeat A"
repeat_region 70..72
/note="direct repeat 1"
repeat_unit 97..103
/note="inverted repeat A"
repeat_region 104..106
/note="direct repeat 1"
CDS 106..1047
/note="Thymidylate synthase (aa 1-313)"
/codon_start=1
/protein_id="CAA26178.1"
/db_xref="GI:37479"
/db_xref="SWISS-PROT:P04818"
/translation="MPVAGSELPRRLPPAQRDARPRPHGLQVLQIOHILRCG
VKDRDTGTLTVFGQARYSLRDEFFLLTKRVFKVGLLELLMFKGSTNAKELS
SKGKIVDANGSRDLDSLGFSTREGDIPVYGFQWRFHGAERYWESDYSGQVQ
LQVDTITKIPDDRIIMCANWPRDLPLMALPCFVYVNSLSQLYQSSGD
MGLGVFNIAISVALLTYMIAHITGLKPGDFIHTGLDAHIYLNIEPLKIQLOREPRPF
PKRLIRKVEKIDFKAEQIEGYNPHPTIMKXAV"
misc_feature 1231..1236
/note="pot. polyadenylation signal"
misc_feature 1519..1524
/note="pot. polyadenylation signal"
polyA_site 1536
/note="polyadenylation site"
BASE COUNT 390 a 369 c 399 g 378 t
ORIGIN
Query Match 25.0%; Score 297; DB 9; Length 1536;
Best Local Similarity 100.0%; Pred. No. 4.8e-157;
Matches 297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 833 CCACCTGGCTGCTCCGTCGCGCGCGCCACCTGGCTGCTCCGTCGCTCCGCGCGCCAC 892
Db 14 CCACCTGGCTGCTCCGTCGCGCGCGCCACCTGGCTGCTCCGTCGCTCCGCGCGCCAC 73
QY 893 TTCGCTGCTCCGTCGCTCCGTCGCGCGCGCCACCTGGCTGCTCCGTCGCTCCGCGCGCCAC 952
Db 74 TTCGCTGCTCCGTCGCTCCGTCGCGCGCGCCACCTGGCTGCTCCGTCGCTCCGCGCGCCAC 133
QY 953 GCCGCGCTTCGCTCCGTCGCGCGCGCCACCTGGCTGCTCCGTCGCTCCGCGCGCCAC 1012
Db 134 GCCGCGCTTCGCTCCGTCGCGCGCGCCACCTGGCTGCTCCGTCGCTCCGCGCGCCAC 193
QY 1013 AGCTGCTGCTGCTCCGTCGCTCCGTCGCGCGCGCCACCTGGCTGCTCCGTCGCTCCGCGCGCCAC 1072
Db 194 AGCTGCTGCTGCTCCGTCGCTCCGTCGCGCGCGCCACCTGGCTGCTCCGTCGCTCCGCGCGCCAC 253
QY 1073 GCACGGGCACCGGACCCCTGCTGGTATTCGGCATGCGAGCGCGCTACAGCTGAGAG 1129
Db 254 GCACGGGCACCGGACCCCTGCTGGTATTCGGCATGCGAGCGCGCTACAGCTGAGAG 310
Search completed: November 29, 2002, 02:33:58
Job time : 3313.34 secs

b 361 ATCTAAAGAGCTGTTTCCAAAGAGAGTGAATCTGGGATGCCAATGGATCCCGAGACT 420
y 455 TTTTGGACAGCTGGGATTTCTCCACGACGAGAGAGGAGCTTTGGGCCCGAGTTATGGCT 514
b 421 TTTTGGACAGCTGGGATTTCTCCACGACGAGAGAGGAGCTTTGGGCCCGAGTTATGGCT 480
y 515 TCCAGTGGAGGCAATTTTGGGCGAGATPACAGAGATATGGAATCAGATTATTCAGGACAGS 574
b 481 TCCAGTGGAGGCAATTTTGGGCGAGATPACAGAGATATGGAATCAGATTATTCAGGACAGS 540
y 575 GAGTTGACCACTGCAAGAGAGTGTGACACATCAAAACCAACCTTGACGACAGAGAA 634
b 541 GAGTTGACCACTGCAAGAGAGTGTGACACATCAAAACCAACCTTGACGACAGAGAA 600
y 635 TCATCATGTGGCTTGAATCCAGAGATCTTCTCTGATGGCGCTGCTCCATGCCATG 694
b 601 TCATCATGTGGCTTGAATCCAGAGATCTTCTCTGATGGCGCTGCTCCATGCCATG 660
y 695 CCTCTGCCAGTCTTATGTGTGAACAGTGAGCTGTCTGCTGCCAGCTGTACAGAGATCGG 754
b 661 CCTCTGCCAGTCTTATGTGTGAACAGTGAGCTGTCTGCTGCCAGCTGTACAGAGATCGG 720
y 755 GAGACATGGGCTCGGTGTGCTTTCAACATCGCCAGCTACGCCCTGCTCACTGATCA 814
b 721 GAGACATGGGCTCGGTGTGCTTTCAACATCGCCAGCTACGCCCTGCTCACTGATCA 780
y 815 TTGCGCATACATCGAGGCTGAAAGCTGAGTGTATATACATCTTTGGGAGATGACATA 874
b 781 TTGCGCATACATCGAGGCTGAAAGCTGAGTGTATATACATCTTTGGGAGATGACATA 840
y 875 TTCTACTGAATCAATCGAGGCTGAAAGCTGAGTGTATATACATCTTTGGGAGATGACATA 934
b 841 TTCTACTGAATCAATCGAGGCTGAAAGCTGAGTGTATATACATCTTTGGGAGATGACATA 900
y 935 CAAAGCTCAGAGATCTTCTGAAAGTGTGAGAAATGATGACTTCAAAGCTGAAAGCTTTC 994
b 901 CAAAGCTCAGAGATCTTCTGAAAGTGTGAGAAATGATGACTTCAAAGCTGAAAGCTTTC 960
y 995 AGATTGAAGGATCAATCCGATCCCACTATTAAATGGAATGGCTTTTGGGCTT 1054
b 961 AGATTGAAGGATCAATCCGATCCCACTATTAAATGGAATGGCTTTTGGGCTT 1020

RESULT 5
L562750/c 1021 bp mRNA linear EST 16-FEB-2001
OCUS
DEFINITION AL562750 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC018YA14 3
prime, mRNA sequence.

CCESION AL562750
ERSION AL562750.1 GI:12911478

KEYWORDS EST.

ORCE human.

ORGANISM Homo sapiens

EuKaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1021)

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: sgrif@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. 1021

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DC018YA14"

/clone_lib="LTI_NFL003_NBC3"

/sex="male"

/tissue_type="neuroblastoma cells"

/lab_host="DH10B"

FEATURES
source

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime and enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 276 a 230 c 203 g 307 t 5 others

ORIGIN

Query Match 61.4%; Score 943.8; DB 9; Length 1021;
Best Local Similarity 97.8%; Pred. No. 5.3e-235;
Matches 962; Conservative 3; Mismatches 18; Indels 1; Gaps 1;

Qy 529 TTTGGGCGAGATACAGAGATATGGAATCAGATTATTTCAGGACAGGAGTTGACCACTG 588

Db 984 TTTGGGCGAGAAAAAGAGATATGAAATCAGATTATTTCAGAACAGGAGTTGACCACTG 925

Qy 589 CAAAGAGTATTGACACCATCAAAACCAACCTTGACGACAGAGATCATCATGTGCGCT 648

Db 924 CAAAGAGTATTGACACCATCAAAACCAACCTTGACGACAGAGATCATCATGTGCGCT 865

Qy 649 TGAAT-CCAGAGATCTTCTCTGATGGCGCTGCTCCATGCCATGCCCTCTGCCAGTT 707

Db 864 TGAATCCAGAGATCTTCTCTGATGGCGCTGCTCCATGCCATGCCCTCTGCCAGTT 805

Qy 708 CTATGTGTTGACAGTGTCTTCCAGCTGTACAGAGATCGGAGACATGGGCT 767

Db 804 CTATGTGTTGACAGTGTCTTCCAGCTGTACAGAGATCGGAGACATGGGCT 745

Qy 768 CGGTGTGCTTTCAATCGCGCAGCTACGCCCTGCTCGTACATGATGCGCACATCAC 827

Db 744 CGGTGTGCTTTCAATCGCGCAGCTACGCCCTGCTCGTACATGATGCGCACATCAC 685

Qy 828 GGGCTTGAAGCCAGTGTCTTATACACATTTGGAGATGCATATTTACCTGAATCA 887

Db 684 GGGCTTGAAGCCAGTGTCTTATACACATTTGGAGATGCATATTTACCTGAATCA 625

Qy 888 CATTGAGCCACTGAAAAATTCAGCTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGGAT 947

Db 624 CATTGAGCCACTGAAAAATTCAGCTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGGAT 565

Qy 948 TCTTCCAAAAGTTGAGAAAATTCATGACCTTCAAAGCTGAAGCTTTTCAGATTGAAGGTA 1007

Db 564 TCTTCCAAAAGTTGAGAAAATTCATGACCTTCAAAGCTGAAGCTTTTCAGATTGAAGGTA 505

Qy 1008 CAATCCGATCCCACTATTAAAAATGGAATGGCTGTTTAGGGTGTCTTCAAAGGAGCTNG 1067

Db 504 CAATCCGATCCCACTATTAAAAATGGAATGGCTGTTTAGGGTGTCTTCAAAGGAGCTNG 445

Qy 1068 AAGGATATTGTCAGTCTTTAGGGGTGGGCTGGATGCGAGGTAAGGTTCTTTTGGCTC 1127

Db 444 AAGGATATTGTCAGTCTTTAGGGGTGGGCTGGATGCGAGGTAAGGTTCTTTTGGCTC 385

Qy 1128 TAAAGAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1187

Db 384 TAAAGAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 325

Qy 1188 GATGTTGCCACTGCGAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1247

Db 324 GATGTTGCCACTGCGAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 265

Qy 1248 TAACTCACTGAGGATCTGACATCTGAGGTATGACAAAGTGAAGGAGATGAAATG 1307

Db 264 TAACTCACTGAGGATCTGACATCTGAGGTATGACAAAGTGAAGGAGATGAAATG 205

Qy 1308 TATGTCCTCTTAGCAAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1367

Db 204 TATGTCCTCTTAGCAAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 145


```

FEATURES
  source
    Location/Qualifiers
      1..935
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="CS001060Y04"
        /clone_lib="LTI_NFL006_PL1"
        /lab_host="DH10B"
        /note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
        cDNA was primed with a NotI-oligo(dT) primer. Five prime
        end enriched, double-stranded cDNA was digested with Not I
        and cloned into the Not I and Eco RV sites of the
        pCMVSPORT 6 vector. Library was constructed by Life
        Technologies. Contact : Feng Liang Life Technologies, a
        division of Invitrogen 9800 Medical Center Drive Rockville
        , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
        fliang@lifetech.com URL :
        http://fulllength.invitrogen.com"
      208 a 279 c 254 g 190 t 4 others
      BASE COUNT
      RIGIN
        Query Match 60.7%; Score 933; DB 9; Length 935;
        Best Local Similarity 99.8%; Pred. No. 3.4e-232;
        Matches 931; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

  Y 30 CGTCCGCGCGCCACTTGGCTCGCTCCGCTCCGCGCCACTTGGCTCGCTCGCTC 89
  b 1 CGTCCGCGCGCCACTTGGCTCGCTCCGCTCCGCGCCACTTGGCTCGCTCGCTC 60
  Y 90 CCCCGCCGCGCGCCACTTGGCTCGCTCCGCTCCGCGCCACTTGGCTCGCTCGCTC 149
  b 61 CCCCGCCGCGCGCCACTTGGCTCGCTCCGCTCCGCGCCACTTGGCTCGCTCGCTC 120
  Y 150 CCGCCGACAGAGCGGAGCGCGCGCTCGCTCCGCTCCGCGCCACTTGGCTCGCTCGCTC 209
  b 121 CCGCCGACAGAGCGGAGCGCGCGCTCGCTCCGCTCCGCGCCACTTGGCTCGCTCGCTC 180
  Y 210 CGAGATCAACACATCTCCGCTCGCGCTCGAGGAGGACACCGCGCGCCACTTGGCTCGCTCGCTC 269
  b 181 CGAGATCCAAACATCTCCGCTCGCGCTCGAGGAGGACACCGCGCGCCACTTGGCTCGCTCGCTC 240
  Y 270 CTGTGCTGATTCGGCATGACGCGCGCTACAGCTGAGAGATGAATTCCTCTGCTGAC 329
  b 241 CTGTGCTGATTCGGCATGACGCGCGCTACAGCTGAGAGATGAATTCCTCTGCTGAC 300
  Y 330 AACCAAGCTGTGTTCTGGAAGGCTGTTTTCGAGAGTGTGCTGTGTTTATCAAGGATC 389
  b 301 AACCAAGCTGTGTTCTGGAAGGCTGTTTTCGAGAGTGTGCTGTGTTTATCAAGGATC 360
  Y 390 CACAAATGCTAAAGAGCTGTCTTCCAAAGGAGTGAATCTGGGATGCCAATGGATCCCG 449
  b 361 CACAAATGCTAAAGAGCTGTCTTCCAAAGGAGTGAATCTGGGATGCCAATGGATCCCG 420
  Y 450 AGACTTTTGGACAGCTCGGATTCCTCCAGAGAGAGGAGCTTGGGCGCCAGTTTA 509
  b 421 AGACTTTTGGACAGCTCGGATTCCTCCAGAGAGAGGAGCTTGGGCGCCAGTTTA 480
  Y 510 TGGCTTCAGTGGAGGCAATTTTGGGCGAGATACAGAGATATGAATTCAGTATTCAGG 569
  b 481 TGGCTTCAGTGGAGGCAATTTTGGGCGAGATACAGAGATATGAATTCAGTATTCAGG 540
  Y 570 ACAGGAGTTCACCACTGCAAGAGTGTATTCACACCATCAAAACCAACCTTCAGCAG 629
  b 541 ACAGGAGTTCACCACTGCAAGAGTGTATTCACACCATCAAAACCAACCTTCAGCAG 600
  Y 630 AAGATCATCATGTGCGCTTGAATCCAAAGAGATCTTCTGTATGCGCTGCTCCATG 689
  b 601 AAGATCATCATGTGCGCTTGAATCCAAAGAGATCTTCTGTATGCGCTGCTCCATG 660
  Y 690 CCATGCCCTCTGCGCAGTTCATGTGTTGAAAGTGTGCTTCGCGAGCTGTACACAG 749
  b 661 CCATGCCCTCTGCGCAGTTCATGTGTTGAAAGTGTGCTTCGCGAGCTGTACACAG 720
  Y 750 ATCGGAGACATGGGCGCTCGGTGTGCTTTCAACATCGCCAGCTACGCCCTGTGCTACGTA 809
  Db 721 ATCGGAGACATGGGCGCTCGGTGTGCTTTCAACATCGCCAGCTACGCCCTGTGCTACGTA 780
  Qy 810 CATGATTCGCACATCAACCGGCTGAGCCAGTGTGCTTATACACACTTTGGAGATGC 869
  Db 781 CATGATTCGCACATCAACCGGCTGAGCCAGTGTGCTTATACACACTTTGGAGATGC 840
  Qy 870 ACATTTTACCTGATCATCATCGACCACTGAAATTCAGCTTCAGCGAGAACCCAGACC 929
  Db 841 ACATTTTACCTGATCATCATCGACCACTGAAATTCAGCTTCAGCGAGAACCCAGACC 900
  Qy 930 TTTCCCAAGCTCAGGATCTTTCGAAAAGTTGAGA 964
  Db 901 TTTCCCAAGCTCAGGATCTTTCGAAAAGTTGAGA 935
  RESULT 8
  AL551990 938 bp mRNA linear EST 16-FEB-2001
  LOCUS AL551990 LTI_NFL006_PL2 Homo sapiens cDNA clone CS001060Y04 5
  DEFINITION prime, mRNA sequence.
  ACCESSION AL551990
  VERSION AL551990.1 GI:12890459
  KEYWORDS EST.
  SOURCE human.
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE 1 (bases 1 to 938)
  AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE Full-length cDNA libraries and normalization
  JOURNAL Unpublished (2001)
  COMMENT Contact: Genoscope
  Genoscope - Centre National de Sequencage
  BP 191 91006 EVRY cedex - France
  Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
  FEATURES
    source
      Location/Qualifiers
        1..938
          /organism="Homo sapiens"
          /db_xref="taxon:9606"
          /clone="CS001060Y04"
          /clone_lib="LTI_NFL006_PL2"
          /tissue_type="placenta"
          /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
          was primed with a NotI-oligo(dT) primer. Five prime end
          enriched, double-stranded cDNA was digested with Not I and
          cloned into the Not I and Eco RV sites of the pCMVSPORT 6
          vector. Library was normalized. Library was constructed by
          Life Technologies. Contact : Feng Liang Life Technologies,
          a division of Invitrogen 9800 Medical Center Drive
          Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
          Email : fliang@lifetech.com URL :
          http://fulllength.invitrogen.com"
        202 a 288 c 259 g 188 t 1 others
        BASE COUNT
        ORIGIN
          Query Match 60.7%; Score 931.6; DB 9; Length 938;
          Best Local Similarity 99.9%; Pred. No. 7.8e-232;
          Matches 931; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

  Qy 14 CCACTTGGCTCGCTCCGCTCCGCGCCACTTGGCTCGCTCGCTCGCGCCAGC 73
  Db 7 CCACTTGGCTCGCTCCGCTCCGCGCCACTTGGCTCGCTCGCTCGCGCCAGC 56
  Qy 74 TTGCGCTGCTCGCTCCGCTCCGCGCCACTTGGCTCGCTCGCTCGCGCCAGC 133
  Db 67 TTGCGCTGCTCGCTCCGCTCCGCGCCACTTGGCTCGCTCGCTCGCGCCAGC 126
  Qy 134 GCGGCGCTTGGCTCCGCTCCGCGCCACTTGGCTCGCTCGCTCGCGCCAGC 193
  Db 127 GCGGCGCTTGGCTCCGCTCCGCGCCACTTGGCTCGCTCGCTCGCGCCAGC 186
  Qy 194 AGCTGAGTACCTGGGCGGAGATCCACATCTCTCGCTGCGGCTCAGGAAGAGCAGC 253

```

```

187 AGCTGACGACCTGCGGAGATCAACATCTCCGTCGGGCTCAGGAGGACACC 246
254 GCACGGGACCGGACACCTGCGGTATTCGGCATCGAGCGCGCTACAGCTCAGAGATG 313
247 GCACGGGACCGGACACCTGCGGTATTCGGCATCGAGCGCGCTACAGCTCAGAGATG 306
314 AATTCCTCTGCTGCACACCAACAGTGTGTCTTGAAGGGTGTTCGAGAGTTCCTGT 373
307 AATTCCTCTGCTGCACACCAACAGTGTGTCTTGAAGGGTGTTCGAGAGTTCCTGT 366
374 GGTTCATCAAGGGATCCACAAATGCTAAAGAGTGTCTTCAAGGGATGAAATCTGGG 433
367 GGTTCATCAAGGGATCCACAAATGCTAAAGAGTGTCTTCAAGGGATGAAATCTGGG 426
434 ATGCCAATGGATCCCGAGACTTTTGGACAGCTGGATTCACACAGAGAGAGGGG 493
427 ATGCCAATGGATCCCGAGACTTTTGGACAGCTGGATTCACACAGAGAGAGGGG 486
494 ACTTGGGCCAGTTTATGGCTTCAGTGGAGGCAATTTGGGCGAGATACAGAGATATGG 553
487 ACTTGGGCCAGTTTATGGCTTCAGTGGAGGCAATTTTGGGCGAGATACAGAGATATGG 546
554 AATCAGATTTTTCAGGACAGGAGTTCACCACTGCAAGAGTGTTCACACCATCAAAA 613
547 AATCAGATTTTTCAGGACAGGAGTTCACCACTGCAAGAGTGTTCACACCATCAAAA 606
614 CCAACCTTCAGACAGAGAGATCATCATGTGCGCTTGAATCCAGAGATCTTCCTGA 673
607 CCAACCTTCAGACAGAGAGATCATCATGTGCGCTTGAATCCAGAGATCTTCCTGA 666
674 TGGCGTGTCTCCATGCGCATCTGCGAGTTCATGTGTGTGTGAAAGTGTGTCT 733
667 TGGCGTGTCTCCATGCGCATCTGCGAGTTCATGTGTGTGAAAGTGTGTCT 726
734 GCCAGTGTACAGAGATCGGGAGACATGGGCTCGGTGTGCTTCAACATGCCAGCT 793
727 GCCAGTGTACAGAGATCGGGAGACATGGGCTCGGTGTGCTTCAACATGCCAGCT 786
794 ACGCTGTCTCAGTACATGATTCGGACATCAGGCGCTGAAGCCAGTGTATATAC 853
787 ACGCTGTCTCAGTACATGATTCGGACATCAGGCGCTGAAGCCAGTGTATATAC 846
854 ACATTTGGAGATGACATATTTACCTGAATCAGATCAGGCGCTGAAGCCAGTGTAT 913
847 ACATTTGGAGATGACATATTTACCTGAATCAGATCAGGCGCTGAAGCCAGTGTAT 906
914 AGCGAGACCCAGACTTTCCCAAGCTCAGG 945
907 AGCGAGACCCAGACTTTCCCAAGCTCAGG 938

```

```

RESULT 9
AL515911 995 bp mRNA linear EST 13-FEB-2001
LOCUS AL515911 LTI_NFL011_NBC1 Homo sapiens cDNA clone CSODA001YD20 5
DEFINITION prime, mRNA sequence.
ACCESSION AL515911
VERSION AL515911.1 GI:12779404
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 995)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
          Genoscope - Centre National de Sequencage
          BP 191 91006 EVRY cedex - France
          Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
          Location/Qualifiers
FEATURES

```

```

1. .995
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODA001YD20"
/clone_lib="LTI_NFL011_NBC1"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library is not normalized, but is the control for
the normalized libraries. Library was constructed by Life
Technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 239 a 272 c 266 g 209 t 9 others
ORIGIN
Query Match 60.2%; Score 924; DB 9; Length 995;
Best Local Similarity 98.1%; Pred. No. 7.6e-230;
Matches 977; Conservative 9; Mismatches 5; Indels 5; Gaps 5;
QY 53 CGCGCGGCACCTTCGCTCCGTCCTCCGCGCGCGCGCGCGCGCGCGCGCGCTC 122
Db 1 CGCGCGGCACCTTCGCTCCGTCCTCCGCGCGCGCGCGCGCGCGCGCGCTC 60
QY 123 GGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 182
Db 61 GGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 120
QY 183 GCCGCAAGGGGAGCTGCAGTACCTGGGGCAGATCCAAACATCTCCGCTCGCGCGCTC 242
Db 121 GCCGCAAGGGGAGCTGCAGTACCTGGGGCAGATCCAAACATCTCCGCTCGCGCGCTC 180
QY 243 GAAGAGCAGCGGACCGGGACCGGGACCGCTGTGCGTATTCGGCATGCGAGCGCGCTCAG 302
Db 181 GAAGAGCAGCGGACCGGGACCGGGACCGCTGTGCGTATTCGGCATGCGAGCGCGCTCAG 240
QY 303 CTTGAGAGATGAATTCCTCTCTGCAACAACAACTGTGTCTCGAGAGGCTGTTTGA 362
Db 241 CCTGARATATGAATTCCTCTCTGCAACAACAACTGTGTCTCGAGAGGCTGTTTGA 300
QY 363 GGAGTCTGTGTTTATCAAGGATCCAAATGCTAAAGAGCTGTCTTCCAAAGGAGT 422
Db 301 GGAGTCTGTGTTTATCAAGGATCCAAATGCTAAAGAGCTGTCTTCCAAAGGAGT 360
QY 423 GAAATCTGGGATGCCAATGGATCCGAGACTTTTGGACAGCTTCTCCACCAG 482
Db 361 AATAATCTGGGATGCCAATGGATCCGAGACTTTTGGACAGCTTCTCCACCAG 420
QY 483 AGAAGAGGAGACTTGGGCGCCAGTTTATGGCTTCCAGTGGAGGCAATTTGGGGCAGAATA 542
Db 421 AGAAGAGGAGACTTGGGCGCCAGTTTATGGCTTCCAGTGGAGGCAATTTGGGGCAGAATA 480
QY 543 CAGAGATATGAATCAGATTATTCAGGACAGGAGTGTGCAACTGCCAAGAGTGTGA 602
Db 481 CAGAGATATGAATCAGATTATTCAGGACAGGAGTGTGCAACTGCCAAGAGTGTGA 540
QY 603 CACCATCAAAACCAACCTTCAGCAGAGAAGATCATCATGTGCGCTTGGAAATCAAGAGA 662
Db 541 CACCATCAAAACCAACCTTCAGCAGAGAAGATCATCATGTGCGCTTGGAAATCAAGAGA 600
QY 663 TCTTCTCTCATGCGCTCCCTCCATGCGCATGCGCTCTGCGAGTCTATGTGTGAACAG 722
Db 601 TCTTCTCTCATGCGCTCCCTCCATGCGCATGCGCTCTGCGAGTCTATGTGTGAACAG 660
QY 723 TGAGCTGTCTGCGAGCTGTACAGAGATCGGAGAGATGGGCTCGGTGTGCTTTCAA 782
Db 661 TGAGCTGTCTGCGAGCTGTACAGAGATCGGAGAGATGGGCTCGGTGTGCTTTCAA 720

```


633	QY	AATCATCATGTGCGCTTGGAAATCCAAAGAGATCTTCTCTGATGGCGCTGCCTCCATGCCA	692	
	Db	601	AATCATCATGTGCGCTTGGAAATCCAAAGAGATCTTCTCTGATGGCGCTGCCTCCATGCCA	660
693	QY	TGCCCTCTGCCAGTTCTTAGTGTGGGAACAGTGAGCTGTCTGCCAGCTGTATCCAGAGATC	752	
	Db	661	TGCCCTCTGCCAGTTCTTAGTGTGGGAACAGTGAGCTGTCTGCCAGCTGTATCCAGAGATC	720
753	QY	GGGAGACATGGGGCTTCGGTGTGGCTTTGAAATGCCAGCTACGCCCTGCTCAGTACAT	812	
	Db	721	GGGAGACATGGGGCTTCGGTGTGGCTTTGAAATGCCAGCTACGCCCTGCTCAGTACAT	780
813	QY	GATTGGCGACATCAGCGGCGCTGAAGCCAGGTGACTTTATACACAC--TTTGGGAGATGCA	870	
	Db	781	GATTGGCGACATCAGCGGCGCTGAAGCCAGGTGACTTTATACACAC--TTTGGGAGATGCA	840
871	QY	CATATTTTACCTCGTAATCACATCGAGCCACTG--AAATTTTCAGCTTTCAGCG--AGAAACCCAGA	927	
	Db	841	CATATTTTACCTCGTAATTCCTATCGAGCCACTTGGAAATTTTCAGCTTCAGCGAGAACCCAGA	900
928	QY	CTTTT--CCCAAAGCTCAGGATTTCTT--CGAAAAAGTTGAGAAAATTTGATGA--CTTCAAG	982	
	Db	901	CTTTTCCCAAAGCTCAGGATTTCTTCCAAAAGGTTGAGAAAATTTGATGACCTTCCAAG	960
983	QY	CTGAAGACTTT--CAGATGGAAGGTACATTCGATCCCACTATTAAATGGAATGGC	1041	
	Db	961	CTGAAGAAATTTTCAGAAATTGAAGGACAATCCCGCATCATCTTTTAAATGGAATGGG	1021
1041	QY	T 1041		
1021	Db	T 1021		

RESULT 14					
ALS15910/c	ALS15910	923 bp	mRNA	linear	EST 13-FEB-2001
LOCUS	ALS15910	LTT NFL011	NBCI Homo sapiens cDNA clone CSODRA001D20_3		
DEFINITION			prime. mRNA sequence.		

SOURCE	ORGANISM	Homosapiens
human.	Human	Human
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	

1 (bases 1 to 923)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)

CONTACT : GENOSCOPE
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
mailto:contact@genoscope.cns.fr Web : www.genoscope.cns.fr

```

FEATURES
source
Location/Qualifiers
1. .923
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="CS0DA001YD20"
/clone_lib="LTI_NFL011_NBC1"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab host="DH10B"
/notes="Organ: brain; Vector: pCMWSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMWSPORT 6
vector. Library is not normalized, but is the control for
the normalized libraries. Library was constructed by life
Technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

```

RESULT 15

AL578646/c
LOCUS
AL578646 LTI NFL006 PL2 Homo sapiens cDNA clone CSODK004YL08 3
prime, mRNA sequence.
AL578646
AL578646.1 GI:12942922
EST.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 906)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 906
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODK004YL08"
/clone_lib="LTI NFL006 PL2"
/tissue_type="placenta"
/note="Vector: pCMWSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMWSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Peng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
ASE COUNT 256 a 200 c 183 g 264 t 3 others
ORIGIN

Query Match 57.8%; Score 887.6; DB 9; Length 906;
Best Local Similarity 99.0%; Pred. No. 2.3e-220;
Matches 898; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

605 CCATCAAAACCAACCTGACGACAGAGAAATCATATGTCGCTTGGATCCAGAGATC 664
|||||
b 906 CCATCAAAACCAACCTGACGACAGAGAAATCATATGTCGCTTGGATCCAGAGATC 847
|||||
y 665 TTCCTCTGATGGCGCTGCTCCATGCCATGCCCTCTGCCAGTTCATATGTGGTGAACAGTG 724
|||||
bb 846 TTCCTCTGATGGCGCTGCTCCATGCCATGCCCTCTGCCAGTTCATATGT-GTGAACAGTG 788
|||||
y 725 AGCTGTCTGCGCAGCTGTACACAGATCGGAGACATGGGCTCGGTGCGCTTTCACCA 784
|||||
bb 787 AGCTGTCTGCGCAGCTGTACACAGATCGGAGACATGGGCTCGGTGCGCTTTCACCA 728
|||||
y 785 TCGCCAGAGTACGCCCTGCTCAGCTACATGATTCGCGACATCAACGGGCTGAAGCCAGTG 844
|||||
bb 727 TCGCCCAATGACGCCCTGCTCAGCTACATGATTCGCGACATCAACGGGCTGAAGCCAGTG 668
|||||
y 845 ACTTTATACACACTTTGGGAGATGCATATTTTACCTGAATTCACATCGAGCCACTGAAA 904
|||||
bb 667 ACTTTATACACACTTTGGGAGATGCATATTTTACCTGAATTCACATCGAGCCACTGAAA 608
|||||
y 905 TTCAGCTTCAGCGAGAACCCAGACTTTCCTCCAAAGCTCAGGATTCCTCGAAAAGTTGAGA 964
|||||
bb 607 TTCAGCTTCAGCGAGAACCCAGACTTTCCTCCAAAGCTCAGGATTCCTCGAAAAGTTGAGA 548
|||||
y 965 AAATTGATGACTTCAAAGCTGAAGACTTTTCAGATTCGAAGGGTACAAATCCGATCCAACTA 1024
|||||
bb 547 AAATTGATGACTTCAAAGCTGAAGACTTTTCAGATTCGAAGGGTACAAATCCGATCCAACTA 488
|||||
y 1025 TTAATAATGGAATGGCTGTTTGGGTGCTTTTCAAGAGAGCTNGAAGGATATTGTCACTCT 1084
|||||

Db 487 TTAATAATGGAATGGCTGTTTGGGTGCTTTCAAAGAGCTTTGAAGGATATTGTCACTCT 428
|||
Qy 1085 TTAGGGGTTGGGCTGATGCCGAGTAAAGTTCTTTTGGCTCTTAAAGANAGGAACT 1144
|||
Db 427 TTAGGGGTTGGGCTGATGCCGAGTAAAGTTCTTTTGGCTCTTAAAGANAGGAACT 368
|||
Qy 1145 AGGTCAAAATCTGTCCGAGACCTATCAGTTATTAATTTTAAAGATGTTGCCACTGGCA 1204
|||
Db 367 AGGTCAAAATCTGTCCGAGACCTATCAGTTATTAATTTTAAAGATGTTGCCACTGGCA 308
|||
Qy 1205 AATCTAACTGTGCCAGTTCTTTTCCATAATAAAGGCTTTGAGTTAACTCACTGAGGGTAT 1264
|||
Db 307 AATCTAACTGTGCCAGTTCTTTTCCATAATAAAGGCTTTGAGTTAACTCACTGAGGGTAT 248
|||
Qy 1265 CTGACAAATGCTGAGGTTATGAACAAACTGAGGAGAAATGAATGTATGTCTTTAGCAAA 1324
|||
Db 247 CTGACAAATGCTGAGGTTATGAACAAAGTGAAGGAGAAATGAATGTATGTCTTTAGCAAA 188
|||
Qy 1325 AACATGTATGTGCATTTCAATCCCACTACTTATAAAGAGGTTGGTGAATTTTCAACAAGC 1384
|||
Db 187 AACATGTATGTGCATTTCAATCCCACTACTTATAAAGAGGTTGGTGAATTTTCAACAAGC 128
|||
Qy 1385 TATTTTGAATATTTTGTAGAAATTTTAAAGATTTTCAAGCTATTCCTCAAAATCTGA 1444
|||
Db 127 TATTTTGGAAATTTTGTAGAAATTTTAAAGATTTTCAAGCTATTCCTCAAAATCTGA 68
|||
Qy 1445 GGGAGCTGAGTAACACCATCGATCATGATGTAGAGTGTGTTTATCAACTTTTANAGTTGTT 1504
|||
Db 67 GGGAGCTGAGTAACACCATCGATCATGATGTAGAGTGTGTTTATCAACTTTTANAGTTGTT 8
|||
Qy 1505 TTATATG 1511
|||
Db 7 TTATATG 1

Search completed: November 28, 2002, 22:40:04

Job time : 2033.8 secs


```

; ; TITLE OF INVENTION: Variants of alternative splicing
; ; FILE REFERENCE: 129181.4 Compugen
; ; CURRENT APPLICATION NUMBER: US/09/724,676
; ; CURRENT FILING DATE: 2000-11-28
; ; NUMBER OF SEQ ID NOS: 97222
; ; SOFTWARE: PatentIn version 3.2
; ; SEQ ID NO 34515
; ; LENGTH: 2042
; ; TYPE: DNA
; ; ORGANISM: Homo sapiens
US-09-724-676-34515

```

```
Query Match          66.7%; Score 1024; DB 5; Length 2042;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1124; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

Qy	1	GGGGGGGGGAGCAACATTGGCTCGCTCCGGCCGGCCAATTGGCTGCCCTCGT	60
Db <td>2042</td> <td>GGGGGGGGGGGAGCAATTGGCTCGCTCCGGCCGGCCAATTGGCTGCCCTCGT</td> <td>1983</td>	2042	GGGGGGGGGGGAGCAATTGGCTCGCTCCGGCCGGCCAATTGGCTGCCCTCGT	1983

Qy 61 CC CG CCG CG CCA C TT CG C CT G C C T C G T C C C G C C G C G C C A T G C C T G T G C C G C 120

ph 182 C C G C C G G A C A T T G C C T G C C T C G T C C C G C C G C C A T G C C T G T G C C G C 1923

Qy 121 TCGAGCTGCGGGCGGGCGCTTGCCCCCGCCGCAAGGACGGGACGCCGAGCCGGCT 180

QY 181 CCGCCGACGGGGACTGCAGTACCTGGGGGAGATCCAAACACATCTTCGCTCGGGGCTC 240

1862 CCGCCGACCGGGGAGCCTCGAAGCTTCCGGGATTCAGGGCGGCTAC 300

QY 241 AGGAAGGACGACCGGACGGGACCGGCACCCCTGTTCGGTATTCAGGGCGGCTAC 300

DB 1802 AGGAAGGACGACCGGACGGGGACCGGACCCCTGTCGGTAATTCCTCTGCTGACACCAACGTTGTTCTCGAAGGGTGTTTTG 360

QY 301 AGCCTGAGAGATGAATTCCTCTGCTGCTGACACCAACGTTGTTCTCGAAGGGTGTTTTG 360

Db 1742 AGCTGAGAGGTGAATCCCTCTGCTGACACCAACGCTGTGTTCTGGAAGGGGTGTTTG 1688

QY 361 GAGGAGTTGCTGTGGTTATCAAGGGGATCCACAAATGCTAAAGAGCTGTTCTTCCAAAGGA 420

Db 1682 GAGGAGTTGCTGTGGTTTATCAAGGGATCCACAAATGCTAAGAGCTGTCTTCCAAAGGA 1623

QY 421 GTGAAAAATCTGGGATGCCAATGGATCCGAGACTTTTTTGGACAGCCTGGGATTTCTCCACC 480

Db 1622 GTGAAAATCTGGGATGCCAATGGATCCGAGACTTTTGGACAGCCTGGGATTTCTCCACC 1563

Dv 481 AGAGAGAGAGAGGAGACTTTGGGGCCAGTTTATGCGTTTCCAGTGAGGCCATTTTGGGGCAGAA 540

Db 1562 AGAGAAGAAGGGGACTTGGGCCAGTTATGGCTTCCAGTGGAGGCCATTTTGGGCAGAA 1503

Db 1502 TACAGATAATGGAATCAGATTATTTCAGGACAGGGAGTTGACCAACTGCGAAGAGTGATT 1443

001
002
003
004
005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053
054
055
056
057
058
059
060
061
062
063
064
065
066
067
068
069
070
071
072
073
074
075
076
077
078
079
080
081
082
083
084
085
086
087
088
089
090
091
092
093
094
095
096
097
098
099
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819

QY 661 GATCTTCTCTGATGGGCTGCTCCATGCCCTCTGCCAGTTCATGTGTGAC 1382

Qy 721 AGTGAGCTGTCCTGCCAGCTGTACAGAGATCGGGAGACATGGGCTCGGTGGCCTTTC 781
=====

Db 1322 AGTGAGCTGTCCTGCCAGCTGTACAGAGATCGGGAGACATGGGCTCGGTGTGCCTTTC 1261

Qy 781 AACATGCCAGGTACGCCCTGCTCACGTACATGATTGGCACATCACGGSCCTGAAGCCA 840

Db 1262 AACATGCCAGGTACGCCCTGCTCACGTACATGATTGGCACATCACGGSCCTGAAGCCA 1202

QY 841 GGTGACTTTATACACACCTTTGGGAGATGCACATATTTTACTGTGATTCATTCGAGGCCACTG 900

Db	1202	GGTGACTTTATACACACTTTGGGAGATGCACATATTTACCTGAATCACATCGAGCACTG	1143
Qy	901	AAAAATTCAGCTTCAGCAGAACCCAGCTCTTCCAAAGCTCAGGATTCCTTCGAAAAAGTT	960
Db	1142	AAAAATTCAGCTTCAGCAGAACCCAGACTTCTCCAAAGCTCAGGATTCCTTCGAAAAAGTT	1083
Qy	961	GAGAAAAATTGATGACTTCAAAGCTGAAGACTTCAGATTCAGAGGTACAAATCCGCATCCA	1020
Db	1082	GAGAAAAATTGATGACTTCAAAGCTGAAGACTTCAGATTCAGAGGTACAAATCCGCATCCA	1023
Qy	1021	ACTATTAAANTGGAATGGCTGTTTAGGCTGCTTCAAAGAGCTNGAAGGATATTGTCA	1080
Db	1022	ACTATTAAANTGGAATGGCTGTTTAGGCTGCTTCAAAGAGCTCGAAGGATATTGTCA	963
Qy	1081	GTCTTTAGGGTTGGCTGGATCCGAGGTAAAAAGTTCTTTTGGCT	1126
Db	962	GTCTTTAGGGTTGGCTGGATCCGAGGTAAAAAGTTCTTTTGGCT	917

```

RESULT 6
US-09-724-676A-34515/c
; Sequence 34515, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34515
; LENGTH: 2042
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-34515

```

Query Match	66.7%;	Score 1024;	DB 5;	Length 2042;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 1124;	Conservative	0;	Mismatches 2;	Indels 0;
Gaps				

QY 1 GGGGGGGGGGACACATTGGCCCTCGCTCCGCGCGGCCACTTGGCTGCCTCGT 60
|||
pb 2042 GGGGGGGGGGACACATTGGCCCTCGCTCCGCGCGGCCACTTGGCTGCCTCGT 1883

QY 61 CCGCGCGCGCACTTCGCTGCCCTCCCGCCCGCGCGGCATGCCCTGTGGCGGCG 120

QY 121 TCGAGCTGCCGCCGCCGCTTGCCCCCGCCGACAGGAGCGGACGCCAGCGCGT 180

QY 181 CCGCGCGACGGGGAGCTGAGTACTCTGGGCAGATCCACACATCTCTCGCTGCGGGGTC 240

1862 CCGCCGACGGGGGAGCTGCAGTACCTCGGCGACGATCTCTAGCAATTCCTCGGCGGCTG 1899

QY 241 AGGAAGGACGACCGGACGGGACACGGGACACCTGTTCGGTATTTCGGCATGTCAGCGCGGTAC 300

Db 1802 AGGAAGGACGACCGGACCGGACCGCTGTTCGGTATTGGCAATCGAGCGCGCTAC 174

QY 301 AGCCTGAGGATGAATTCCCTCTGCTGCAACCAACCGTGTCTTCTGGAAGGGTGTTTTG 360

Db	1742	AGCCTGAGAGGTGAATCCCTCTGCTGACCAACCAACGTCGTGTTCTTGGGAAGGGTGTTTG	1683
QY	361	GAGGAGTTGCTGTGGTTTATCAAGGGATCCACAAATGCTAAAGAGCTGTCTTCCAAAGGA	420

Db 1682 GAGGAGTTGCTGTGGTTTATCAAGGATCCACAATGCTAAAGAGCTGTCTTCCAAGGA 1623
 |||||
 QY 421 GTGAAAAATCTGGGATGCCAATGGATCCCGAGACTTTTGGACAGCCTGGGATTTCTCCACC 480

Db 1622 GTGAAAATCTGGGATGCCAATGGATCCCGAGACITTTTGGACAGCCTGGGATTCCTCCACC 1563

Db 1462 AGAGAGAGAGGAGCTTTGGGCCCGCAGTTATGCTTCCAGTGGAGGCAATTTTGGGGCAGAA 1403
QY 541 TACAGAGATGATGAATCAGATATTTACAGACAGAGGAGTTGACCAACTGCAAGAGAGTGAAT 600
Db 1402 TACAGAGATGATGAATCAGATATTTACAGACAGAGGAGTTGACCAACTGCAAGAGTGAAT 1343
QY 601 GACACCATCAAAACCAACCTGACACAGAGAGATCAATCATATGTCCTTGGGAATCCAGA 660
Db 1342 GACACCATCAAAACCAACCTGACACAGAGAGATCAATCATATGTCCTTGGGAATCCAGA 1283
QY 661 GATCTTCTCTGATGGCGTGCCTCCATGCCATGCTTCCAGAGTTCATATGTTGATGAC 720
Db 1282 GATCTTCTCTGATGGCGTGCCTCCATGCCATGCTTCCAGAGTTCATATGTTGATGAC 1223
QY 721 AGTAGAGCTGCTGCGACAGCTGACAGAGATCGGAGACATGGGCTCGGTGTCCTTTC 780
Db 1222 AGTAGAGCTGCTGCGACAGCTGACAGAGATCGGAGACATGGGCTCGGTGTCCTTTC 1163
QY 781 AACATCGCAGCTACGCCCTGCTCAGTACATGATGGACATCACGGGCTGGAAGCCA 840
Db 1162 AACATCGCAGCTACGCCCTGCTCAGTACATGATGGACATCACGGGCTGGAAGCCA 1103
QY 841 GGTGACTTTATACACATTTGGGAGATGCACATATTTACCTGAATCACATCGAGCCACTG 900
Db 1102 GGTGACTTTATACACATTTGGGAGATGCACATATTTACCTGAATCACATCGAGCCACTG 1043
QY 901 AAAATTCAGCTTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGGATCTTCGAAAAGTT 960
Db 1042 AAAATTCAGCTTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGGATCTTCGAAAAGTT 983
QY 961 GAGAAAATTTGATGACTTCAAAAGCTGAAGACTTTTCAAGGCTGAAAGGTGACATCCGATCCA 1020
Db 982 GAGAAAATTTGATGACTTCAAAAGCTGAAGACTTTTCAAGGCTGAAAGGTGACATCCGATCCA 923
QY 1021 ACTATTAATTTGAAATGGCTTTAGGCTGCTTTCAAGGCTGAAAGGTGACATCCGATCCA 1080
Db 922 ACTATTAATTTGAAATGGCTTTAGGCTGCTTTCAAGGCTGAAAGGTGACATCCGATCCA 863
QY 1081 GTCCTTTAGGGGTTGGGCTGGATGCCGAGGTAAAGTTCTTTTGGCT 1126
Db 862 GTCCTTTAGGGGTTGGGCTGGATGCCGAGGTAAAGTTCTTTTGGCT 817

RESULT 4
US-09-724-676A-34523/c
; Sequence 34523, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-34523
Query Match 66.7%; Score 1024; DB 5; Length 1942;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1124; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCTCCGCGCCGACCTTGGCTCCCTCCGT 60
Db 1942 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCTCCGCGCCGACCTTGGCTCCCTCCGT 1883
QY 61 CCGCGCGCGCCACTTGGCTGCTCCGTCGCTCCGCGCCGCGCCGACCTTGGCTCCCTCCGT 120
Db 1882 CCGCGCGCGCCACTTGGCTGCTCCGTCGCTCCGCGCCGCGCCGACCTTGGCTCCCTCCGT 1823
QY 121 TCGGAGCTGCCGCGCGCCCTTTGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 180

Db 1822 TCGGAGCTGCCGCGCGCCCTTTGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1763
QY 181 CCGCGCGCGCGGAGCTGAGTACCTTGGGGCAGATCCACACATCTCGCTCGGCTGGGGGCTC 240
Db 1762 CCGCGCGCGCGGAGCTGAGTACCTTGGGGCAGATCCACACATCTCGCTCGGCTGGGGGCTC 1703
QY 241 AGGAAGACGACCGCAGCGGCGACCGGCGACCTGTCGGTATTTGGCATGCGAGCGCGCTAC 300
Db 1702 AGGAAGACGACCGCAGCGGCGACCGGCGACCTGTCGGTATTTGGCATGCGAGCGCGCTAC 1643
QY 301 ASCCTGAGAGATGAATTTCCCTCTGCTGCAACAACCAACCTGCTTCTGGAAGGGTGTTTG 360
Db 1642 AGCTTGTAGAGGTGAATTTCCCTCTGCTGCAACAACCAACCTGCTTCTGGAAGGGTGTTTG 1583
QY 361 GAGGAGTTGCTGTGTTTATCAAGGGATCCCAAAATCTAAAGAGCTGTCTCCAGGGA 420
Db 1582 GAGGAGTTGCTGTGTTTATCAAGGGATCCCAAAATCTAAAGAGCTGTCTCCAGGGA 1523
QY 421 GTGAAATCTGGGATGCCAATGGATCCCGAGATCTTTTGGACAGCCCTGGGATTTCCACC 480
Db 1522 GTGAAATCTGGGATGCCAATGGATCCCGAGATCTTTTGGACAGCCCTGGGATTTCCACC 1463
QY 481 AGAGAAGAGGGGACTTTGGGCGCCCTGTTTATGCTTCCAGTGGAGGCAATTTTGGGCGAGAA 540
Db 1462 AGAGAAGAGGGGACTTTGGGCGCCCTGTTTATGCTTCCAGTGGAGGCAATTTTGGGCGAGAA 1403
QY 541 TACAGAGATGAGGATCAGATATTTACAGACAGAGGAGTTGACCAACTGCAAGAGAGTGAAT 600
Db 1402 TACAGAGATGAGGATCAGATATTTACAGACAGAGGAGTTGACCAACTGCAAGAGAGTGAAT 1343
QY 601 GACACCATCAAAACCAACCTGACACAGAGAGATCAATCATGTCGCTTGGGAATCCAGA 660
Db 1342 GACACCATCAAAACCAACCTGACACAGAGAGATCAATCATGTCGCTTGGGAATCCAGA 1283
QY 661 GATCTTCTCTGATGGCGTGCCTCCATGCCATGCTTCCAGAGTTCATATGTTGATGAC 720
Db 1282 GATCTTCTCTGATGGCGTGCCTCCATGCCATGCTTCCAGAGTTCATATGTTGATGAC 1223
QY 721 AGTAGAGCTGCTGCGACAGCTGACAGAGATCGGAGACATGGGCTCGGTGTCCTTTC 780
Db 1222 AGTAGAGCTGCTGCGACAGCTGACAGAGATCGGAGACATGGGCTCGGTGTCCTTTC 1163
QY 781 AACATCGCAGCTACGCCCTGCTCAGTACATGATGGACATCACGGGCTGGAAGCCA 840
Db 1162 AACATCGCAGCTACGCCCTGCTCAGTACATGATGGACATCACGGGCTGGAAGCCA 1103
QY 841 GGTGACTTTATACACATTTGGGAGATGCACATATTTACCTGAATCACATCGAGCCACTG 900
Db 1102 GGTGACTTTATACACATTTGGGAGATGCACATATTTACCTGAATCACATCGAGCCACTG 1043
QY 901 AAAATTCAGCTTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGGATTTCTCGAAAAGTT 960
Db 1042 AAAATTCAGCTTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGGATTTCTCGAAAAGTT 983
QY 961 GAGAAAATTTGATGACTTCAAAAGCTGAAGACTTTTCAAGGCTGAAAGGTGACATCCGATCCA 1020
Db 982 GAGAAAATTTGATGACTTCAAAAGCTGAAGACTTTTCAAGGCTGAAAGGTGACATCCGATCCA 923
QY 1021 ACTATTAATTTGAAATGGCTTTAGGCTGCTTTCAAGGCTGAAAGGTGACATTTGTCA 1080
Db 922 ACTATTAATTTGAAATGGCTTTAGGCTGCTTTCAAGGCTGAAAGGTGACATTTGTCA 863
QY 1081 GTCCTTTAGGGGTTGGGCTGGATGCCGAGGTAAAGTTCTTTTGGCT 1126
Db 862 GTCCTTTAGGGGTTGGGCTGGATGCCGAGGTAAAGTTCTTTTGGCT 817

RESULT 5
US-09-724-676-34515/c
; Sequence 34515, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD

D _b	148	GCCGAGCGCGCTCCGCGCAGCGGGAGCTGCAGTACCTGCGGGGAGATCCACACATCTCTC	207
Q _y	229	CGCTCGGCGCTCAGAAAGGACGACGCGCAGCGCACCGGACCGCTCTCGGTATTCGGCATG	288
D _b	208	CGCTCGGCGCTCAGAAAGGACGACGCGCAGCGCACCGGACCGCTCTCGGTATTCGGCATG	267
Q _y	289	CAGGCGGCTACGCTCAGAGATGAATTCCTCTCTGTCGACCAACAAAGTGTGTCTCG	348
D _b	268	CAGGCGGCTACGCTCAGAGATGAATTCCTCTCTGTCGACCAACAAAGTGTGTCTCG	327
Q _y	349	AAGGCTGTCTTGGAGGATTTGCTGTGTTATTAAGGGAATCAAAATCTTAAGAGCTG	408
D _b	328	AAGGCTGTCTTGGAGGATTTGCTGTGTTATTAAGGGAATCAAAATCTTAAGAGCTG	387
Q _y	409	TCCTTCAAGGAGTGAAATCTCGGATGCCAATGGATCCCGAGACTTTTTCGACAGCTG	468
D _b	388	TCCTTCAAGGAGTGAAATCTCGGATGCCAATGGATCCCGAGACTTTTTCGACAGCTG	447
Q _y	469	GGATTTCCACGAGAGAGGAGCTTGGGCCAGTTATGTTATGCTTCAGTGGAGCAT	528
D _b	448	GGATTTCCACGAGAGAGGAGCTTGGGCCAGTTATGTTATGCTTCAGTGGAGCAT	507
Q _y	529	TTTGGGGCAGATACAGAGATATGGAATCAGATATTACGACAGGGAGTTGACAACTG	588
D _b	508	TTTGGGGCAGATACAGAGATATGGAATCAGATATTACGATATTACGACAGGGAGTTGACAACTG	567
Q _y	589	CAAAAGATGATTGACACCATCAAAACCAAACCCGTGACGACAGAGATCATCATGTGCGCT	648
D _b	568	CABAAGATGATTGACACCATCAAAACCAAACCCGTGACGACAGAGATCATCATGTGCGCT	627
Q _y	649	TGGAATCCAAGAGATCTTCTCTGATGGCGCTGCTCCATGCCATGCCCTCTGCGAGTTT	708
D _b	628	TGGAATCCAAGAGATCTTCTCTGATGGCGCTGCTCCATGCCATGCCCTCTGCGAGTTT	687
Q _y	709	TATGTGGTGAACAGTGAAGCTGCTCTGCGAGCTGTACGAGATCGGGAGACATGGGCTC	768
D _b	688	TATGTGGTGAACAGTGAAGCTGCTCTGCGAGCTGTACGAGAGATCGGGAGACATGGGCTC	747
Q _y	769	GGTGTGCTTTCAACATCGCCAGCTACGCCCTGCTCAGCTACATGATTCGGCACATCAAG	828
D _b	748	GGTGTGCTTTCAACATCGCCAGCTACGCCCTGCTCAGCTACATGATTCGGCACATCAAG	807
Q _y	829	GGCTGAGCGCAGTGACTTTATACACACTTTGGGAGATGCAATATTTCAGTCAATCAC	888
D _b	808	GGCTGAGCGCAGTGACTTTATACACACTTTGGGAGATGCAATATTTCAGTCAATCAC	867
Q _y	889	ATCGAGCCACTGAAATTCAGCTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGGATT	948
D _b	868	ATCGAGCCACTGAAATTCAGCTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGGATT	927
Q _y	949	CTTCGAAAGTTGAGAAATTAATGATCTTCAAAGCTGAAGCTTTGATGTAAGGGTAC	1008
D _b	928	CTTCGAAAGTTGAGAAATTAATGATCTTCAAAGCTGAAGCTTTGATGTAAGGGTAC	987
Q _y	1009	AATCCGATCCAACTATTAAATTCGAAATGGCTGTATTAGGTGCTTTCAAAGGACTNGA	1068
D _b	988	AATCCGATCCAACTATTAAATTCGAAATGGCTGTATTAGGTGCTTTCAAAGGACTNGA	1047
Q _y	1069	AGGATATTGTCAGTCTTTAGGGGTTGGGCTGGATCCGAGCTAAAGTCTTTTTCGTCT	1128
D _b	1048	AGGATATTGTCAGTCTTTAGGGGTTGGGCTGGATCCGAGCTAAAGTCTTTTTCGTCT	1107
Q _y	1129	AAAGAAATAAGAACTAGCTCAAAAATCTCTCCGTGACCTATCATGTTATTAATTTTAAG	1188
D _b	1108	AAAGAAATAAGAACTAGCTCAAAAATCTCTCCGTGACCTATCATGTTATTAATTTTAAG	1167
Q _y	1189	GATGTGTCACCTGGCAAAATGTAATGTGCCAGTTCCTTCGATTAATAAGGCTTTGAGTT	1248
D _b	1168	GATGTGTCACCTGGCAAAATGTAATGTGCCAGTTCCTTCGATTAATAAGGCTTTGAGTT	1227
Q _y	1249	AACTCAGTGAAGGATCTCAGATGCTGAGGTTATGAAACAAAGTCAGGAGATGAATGT	1308
D _b	1228	AACTCAGTGAAGGATCTCAGATGCTGAGGTTATGAAACAAAGTCAGGAGATGAATGT	1287

Qy	1309	ATGTGCTCTTAGCAAAACATGTATGTGCATTTCAATCCACAGCTATTATATAAGAGGTT	13368
Db	1288	ATGTGCTCTTAGCAAAACATGTATGTGCATTTCAATCCACAGCTATTATATAAGAGGTT	1347
Qy	1369	GTTGAATTTTCAACAAGCTATTTTTTGGAAATATTTTTAGAAATATTTTAAAGATTTTCAACAAGCT	1428
Db	1348	GTTGAATTTTCAACAAGCTATTTTTTGGAAATATTTTTAGAAATATTTTAAAGATTTTCAACAAGCT	1407
Qy	1429	ATTCCCTCAAATCTGAGGGAGCTGAGTAACACCATCATCATCATCTAGAGTGTGGTTAT	1488
Db	1408	ATTCCCTCAAATCTGAGGGAGCTGAGTAACACCATCATCATCTAGAGTGTGGTTAT	1467
Qy	1489	GAACTTTA	1496
Db	1468	GAACTTTA	1475

RESULT 3

```

US-09-724-676-34523/c
; Sequence 34523, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34523
; LENGTH: 1942
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-34523

```

Query Match	66.7%	Score 1024:	DB 5:	Length 1942;
-------------	-------	-------------	-------	--------------

Query Match	88.7%	Score 1124, SS 5, Length 1124
Best Local Similarity	99.8%	Pred. No. 0;
Matches 1124; Conservative	0;	Mismatches 2; Indels 0; Gaps 0;

QY	1	GGGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCGCGCCACTTGGCTGCTCCGCTCGT	60
Db	1942	GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCGCGCCACTTGGCTGCTCCGCTCGT	1883
QY	61	CCCGCGCGCCCACTTCGCTGCTCCGTCGCCCGCGCGCGCCATGCTGTGCGCGCG	120
Db	1882	CCCGCGCGCCCACTTTCGCTGCTCCGTCGCCCGCGCGCGCCATGCTGTGCGCGCG	1823
QY	121	TCGGAGCTCCGCGCGCGGCTTGTCCCGCGCGCGCGCGCGAGCGCGCGCT	180
Db	1822	TCGGAGCTCCGCGCGCGGCTTGTCCCGCGCGCGCGCGAGCGCGCGCT	1763
QY	181	CCCGCGCAGCGGGAGCTGCAGTAGCTACCTGGGCGAGATCAAAACACATCTCTCGCTGCGCGCTC	240
Db	1762	CCCGCGCAGCGGGAGCTGCAGTAGCTACCTGGGCGAGATCAAAACACATCTCTCGCTGCGCGCTC	1703
QY	241	AGGAAGGACGACCGCAGCGGCAACGGCACCCCTGTGCTATTTCGGCATACGAGCGCGCTTAC	300
Db	1702	AGGAAGGACGACCGCAGCGGCAACGGCACCCCTGTGCTATTTCGGCATACGAGCGCGCTTAC	1643
QY	301	AGCTGAGAGATGAATTCCTCTGCTGACACCAACGCTGTCTCTGGAGAGGTCGTTTG	360
Db	1642	AGCTGAGAGGTGAATTCCTCTGCTGACCAACGCTGTCTCTGGAGAGGTCGTTTG	1583
QY	361	GAGGAGTTCCTGCTGTTTATCAAGGATCCAAATGCTAAAGAGCTGTCTCCAAAGGA	420
Db	1582	GAGGAGTTCCTGCTGTTTATCAAGGATCCAAATGCTAAAGAGCTGTCTCCAAAGGA	1523
QY	421	GTGAAATCTGGGATGCCAATGGATCCGAGACTTTTTTGGACAGCCTGGGATTCCTCAAC	480
Db	1522	GTGAAATCTGGGATGCCAATGGATCCGAGACTTTTTTGGACAGCCTGGGATTCCTCAAC	1463
QY	481	AGAGAAGAAAGGGAATTGGGCCCAAGTTTATGCTTCAGTGGAGGCAATTTTCGGGCAGAA	540

Db 61 CCAGCGGCGCACTTCCGCTGCTCCGTCGCCGCGCGCGCCATGCTGTGGCGCGC 120
Qy 121 TCGAGAGTCCGCGCGCGCTTGGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 TCGAGAGTCCGCGCGCGCGCTTGGCCCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Qy 181 CCAGCGGCGCGGAGTCCGAGTACCTGGGCGAGATCCAAACACATCTCCGCTGGCGGCTC 240
Db 181 CCAGCGGCGCGGAGTCCGAGTACCTGGGCGAGATCCAAACACATCTCCGCTGGCGGCTC 240
Qy 241 AGAAGGACGACCG 300
Db 241 AGAAGGACGACCG 300
Qy 301 AGCCTGAGAGATGAATTCCTCTGCTGACAAACAAAGCTGTGTTCGAGAGGTGTTT 360
Db 301 AGCCTGAGAGATGAATTCCTCTGCTGACAAACAAAGCTGTGTTCGAGAGGTGTTT 360
Qy 361 GAGAGTGTCTGTGTTTATCAAGGATCCCAAAATCTTAAGAGCTGTCTTCCAGGGA 420
Db 361 GAGAGTGTCTGTGTTTATCAAGGATCCCAAAATCTTAAGAGCTGTCTTCCAGGGA 420
Qy 421 GTGAATCTGGATGCCAATGATCCCGAGACTTTTGGACAGCTTGGGATTTCCACC 480
Db 421 GTGAATCTGGATGCCAATGATCCCGAGACTTTTGGACAGCTTGGGATTTCCACC 480
Qy 481 AGAGAGAGGAGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 481 AGAGAGAGGAGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Qy 541 TACAGAGATGGAATCAGATTAATTCAGGACAGGAGTTCACCACTGCAAGAGTGT 600
Db 541 TACAGAGATGGAATCAGATTAATTCAGGACAGGAGTTCACCACTGCAAGAGTGT 600
Qy 601 GACACATCAAAACCAACCTGACGACAGAGATCATATGTCGCTGGATTCACAGA 660
Db 601 GACACATCAAAACCAACCTGACGACAGAGATCATATGTCGCTGGATTCACAGA 660
Qy 661 GATCTTCTGTGAGGCTGCTCCATGCCATGCTTCCAGCTTCTATGTGTGAAC 720
Db 661 GATCTTCTGTGAGGCTGCTCCATGCCATGCTTCCAGCTTCTATGTGTGAAC 720
Qy 721 AGTGAAGTCTGCTGAGGCTGCTCCATGCCATGCTTCCAGCTTCTATGTGTGAAC 780
Db 721 AGTGAAGTCTGCTGAGGCTGCTCCATGCCATGCTTCCAGCTTCTATGTGTGAAC 780
Qy 781 AACATGCGCAGCTACGCGCTGCTGCTACATGATGTCGACATCAGCGGCTGAGCCA 840
Db 781 AACATGCGCAGCTACGCGCTGCTGCTACATGATGTCGACATCAGCGGCTGAGCCA 840
Qy 841 GGTGATTTATPACACTTTGGGAGATGACATATTTACCTGAATCAGATCAGGACCTG 900
Db 841 GGTGATTTATPACACTTTGGGAGATGACATATTTACCTGAATCAGATCAGGACCTG 900
Qy 901 AAAATTCAGCTTACGAGAACCCAGACCTTCCAAAGCTCAGATTTCTCGAAAGTT 960
Db 901 AAAATTCAGCTTACGAGAACCCAGACCTTCCAAAGCTCAGATTTCTCGAAAGTT 960
Qy 961 GAGAAATTTGATGATTTCAAGCTGAAGATTTTCAAGGAGTCAATCCGATCCA 1020
Db 961 GAGAAATTTGATGATTTCAAGCTGAAGATTTTCAAGGAGTCAATCCGATCCA 1020
Qy 1021 ACTATTAATTTGAAATGCTGATCCGAGGTAAAGTCTTTTCAAGGAGTCAATTTGCA 1080
Db 1021 ACTATTAATTTGAAATGCTGATCCGAGGTAAAGTCTTTTCAAGGAGTCAATTTGCA 1080
Qy 1081 GTCCTTTAGGGGTGGGCTGATCCGAGGTAAAGTCTTTTCAAGGAGTCAATTTGCA 1140
Db 1081 GTCCTTTAGGGGTGGGCTGATCCGAGGTAAAGTCTTTTCAAGGAGTCAATTTGCA 1140
Qy 1141 AACTAGGTCAAAATCTGCTGCTGATCCTATCAGTTATTAATTTTAAAGTGTGCACT 1200
Db 1141 AACTAGGTCAAAATCTGCTGCTGATCCTATCAGTTATTAATTTTAAAGTGTGCACT 1200

Qy 1201 GGCAATGTAATGTGCGAGTCTTCCATAATAAAGGCTTTGAGTTAACTCACTGAGG 1260
Db 1201 GGCAATGTAATGTGCGAGTCTTCCATAATAAAGGCTTTGAGTTAACTCACTGAGG 1260
Qy 1261 GTATCTGACATGCTGAGGTTATGAACAAAGTGGAGGAAATGAATGTATGCTCTTAG 1320
Db 1261 GTATCTGACATGCTGAGGTTATGAACAAAGTGGAGGAAATGAATGTATGCTCTTAG 1320
Qy 1321 CAAAAACATGATGTGCAATTTCAATCCCACTTATAAAGAGTTGGTGAATTTTAC 1380
Db 1321 CAAAAACATGATGTGCAATTTCAATCCCACTTATAAAGAGTTGGTGAATTTTAC 1380
Qy 1381 AAGCTATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 1440
Db 1381 AAGCTATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 1440
Qy 1441 CTGAGGAGCTGAGTGAACCACTCATGATGATGAGTGTGTTTATGAACTTTANAGT 1500
Db 1441 CTGAGGAGCTGAGTGAACCACTCATGATGATGAGTGTGTTTATGAACTTTANAGT 1500
Qy 1501 TGTTTTATGTTGCTATATATAAAGAGTGTCTTGC 1536
Db 1501 TGTTTTATGTTGCTATATATAAAGAGTGTCTTGC 1536

RESULT 2
US-10-240-965-200
; Sequence 200, Application US/10240965
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: SHIFFMAN, Dov
; APPLICANT: SOMOGYI, Roland
; APPLICANT: LAWN, Richard M.
; APPLICANT: SEILHAMER, Jeffrey J.
; APPLICANT: PORTER, Gordon J.
; APPLICANT: MIKITA, Thomas
; APPLICANT: TAL, Julie
; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
; FILE REFERENCE: PA-0025 PCT
; CURRENT APPLICATION NUMBER: US/10/240,965
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/195,106
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 276
; SOFTWARE: PERL Program
; SEQ ID NO 200
; LENGTH: 3298
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 247220.15
; LOCATION: 17, 1531, 1740-1765
; OTHER INFORMATION: a, t, c, g, or other
US-10-240-965-200

Query Match 87.6%; Score 1346; DB 6; Length 3298;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1446; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 49 GCCTGCTCCGTCGCCGCGCGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 108
Db 28 GCCTGCTCCGTCGCCGCGCGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 87
Qy 109 CTTGTGCGGCTGCGAGCTCCGCGCGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCG 168
Db 88 CTTGTGCGGCTGCGAGCTCCGCGCGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCG 147
Qy 169 GCCGAGCGGCTCCGCGCGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 228

Db 361 CGACAGAGATCATGTGCGCTTGGAAATCCAAGAGATCTTCTCTGATGGCGCTGCC 420

Qy 684 TCCATGCCATGCCCTCTGCCAGTCTCTATGTGTAACAGTGTGCTGTCTGCCAGCTGTA 743

Db 421 TCCATGCCATGCCCTCTGCCAGTCTCTATGTGTAACAGTGTGCTGTCTGCCAGCTGTA 480

Qy 744 CCAGAGATCGGAGACATGGGCTCGGCTGTGCTTTCAACATCGCCAGCTTACGCCCTGCT 803

Db 481 CCAGAGATCGGAGACATGGGCTCGGCTGTGCTTTCAACATCGCCAGCTTACGCCCTGCT 540

Qy 804 CAGTACATGATTTGCGACATCATCGGCGCTGGAAGCCAGTCACTTTATATACACTTTGGG 863

Db 541 CAGTACATGATTTGCGACATCATCGGCGCTGGAAGCCAGTCACTTTATATACACTTTGGG 600

Qy 864 AGATGCACATATTTACCTGAAATCAATCGAGCCACTGAAATTCAGTTCAGCGAGAAC 923

Db 601 AGATGCACATATTTACCTGAAATCAATCGAGCCACTGAAATTCAGTTCAGCGAGAAC 660

Qy 924 CAGACTTTCCCAAGCTCAGATTTCTCGAAAGTTGAGAAAT 968

Db 661 CAGACTTTCCCAAGCTCAGATTTCTCGAAAGTTGAGAAAT 705

RESULT 15

BQ057866

LOCUS

AGENCOURT 7047159 NIH_MGC_99 Homo sapiens cdna clone IMAGE:5813622

5', mRNA sequence.

ACCESSION

BQ057866.1 GI:19817206

VERSION

EST.

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 1071)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cdna Library Preparation: Ruben Laboratory
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILNL at:
<http://image.llnl.gov>
Plate: LLCW2065 row: p column: 07
High quality sequence stop: 647.

FEATURES

source

1..1071

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5813622"

/clone_lib="NIH_MGC_99"

/tissue_type="lymphoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: lymph, Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cdna made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT

253 a 300 c 291 g 226 t 1 others

ORIGIN

Query Match 45.3%; Score 696; DB 14; Length 1071;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 TCCGCGCGGCACCTTCCCTGCTCCCTCCCGCGCCGCCATCCTGTGTGCCCG 119

Db 1 TCCGCGCGGCACCTTCCCTGCTCCCTCCCGCGCCGCCATCCTGTGTGCCCG 60

Qy 120 CTCGGAGTTCGCGCGCGCGCCCTTTCGCCCGCGCCGCCATCCTGTGTGCCCG 179

Db 61 CTCGGAGTTCGCGCGCGCGCCCTTTCGCCCGCGCCGCCATCCTGTGTGCCCG 120

Qy 180 TCCGCGCGCACGGGAGCTGAGTACCTGGGGGAGATCCAAACATCTCTCGCTGGGGGT 239

Db 121 TCCGCGCGCACGGGAGCTGAGTACCTGGGGGAGATCCAAACATCTCTCGCTGGGGGT 180

Qy 240 CAGGAAGAGACGACCGACCGGCGACCGGCGCTGTTCGGGTATTCGGCATGCAGCGCGCTA 299

Db 181 CAGGAAGAGACGACCGACCGGCGACCGGCGCTGTTCGGGTATTCGGCATGCAGCGCGCTA 240

Qy 300 CAGCCTGAGAGATGAATTCCTCTGCTGACAAACAAAGCTGTCTTGGAAAGGGTGT 359

Db 241 CAGCCTGAGAGATGAATTCCTCTGCTGACAAACAAAGCTGTCTTGGAAAGGGTGT 300

Qy 360 GGAGGAGTTGCTGTGTGTTATCAAGGAGTCCAAATGCTTAAAGAGCTGTCTTCAAGGG 419

Db 301 GGAGGAGTTGCTGTGTGTTATCAAGGAGTCCAAATGCTTAAAGAGCTGTCTTCAAGGG 360

Qy 420 AGTGAATAATCTGGGATGCCAATGGATCCCGAGACTTTTGGACAGCCTGGGATCTCCAC 479

Db 361 AGTGAATAATCTGGGATGCCAATGGATCCCGAGACTTTTGGACAGCCTGGGATCTCCAC 420

Qy 480 CAGAGAAGAGGGGACTTGGGCCAGTTTATGGCTTCCAGTGGAGGCAATTTGGGGCAGA 539

Db 421 CAGAGAAGAGGGGACTTGGGCCAGTTTATGGCTTCCAGTGGAGGCAATTTGGGGCAGA 480

Qy 540 ATACAGAGATATGGATCAGATTATTCAGGACAGGGAGTTGACCACTGCAAAAGAGTGTAT 599

Db 481 ATACAGAGATATGGATCAGATTATTCAGGACAGGGAGTTGACCACTGCAAAAGAGTGTAT 540

Qy 600 TGACACCATCAAAACCAACCCCTGACGACAGAGAATCATCATGTGCGCTTGGAAATCCAG 659

Db 541 TGACACCATCAAAACCAACCCCTGACGACAGAGAATCATCATGTGCGCTTGGAAATCCAG 600

Qy 660 AGATCTTCTCTGATGGGCTGCGCTCCATGCCATGCCCTCTCCAGTTCTATGTGTGAA 719

Db 601 AGATCTTCTCTGATGGGCTGCGCTCCATGCCATGCCCTCTCCAGTTCTATGTGTGAA 660

Qy 720 CAGTGAAGCTCTCTGCGAGCTGTACCAAGAGATCGGG 755

Db 661 CAGTGAAGCTCTCTGCGAGCTGTACCAAGAGATCGGG 696

Search completed: November 29, 2002, 03:27:57

Job time : 2025.68 secs


```

QY 751 TCGGAGACATGGCC 766
Db 721 TCGGAGACATGGCC 736

RESULT 12
AL542409 935 bp mRNA linear EST 16-FEB-2001
DEFINITION AL542409 LTI FL002_P1 Homo sapiens cDNA clone CS00E010H04 5 prime
, mRNA sequence.
ACCESSION AL542409
VERSION AL542409.1 GI:12874425
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 935)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..935
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS00E010H04"
/clone_lib="LTI FL002_P1"
/lab_host="DH10B"
/notes="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by Life
Technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive, Rockville
, Maryland 20850, USA Fax : (1) 301 610 8371 Email :
filiang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 208 a 279 c 254 g 190 t 4 others
ORIGIN
Query Match 47.6%; Score 731; DB 9; Length 935;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 931; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 30 CCGTCCCGCGCGCCACTTGGCTCGCTCCGTCGCCCGCGCCACTTGGCTCGCTCGCTC 89
Db 1 CCGTCCCGCGCGCCACTTGGCTCGCTCCGTCGCCCGCGCCACTTGGCTCGCTCGCTC 60

QY 90 CCCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 149
Db 61 CCCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120

QY 150 CCGCGCAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 209
Db 121 CCGCGCAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180

QY 210 GCAGATCCAAACATCTCCGCTCGCGCTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 269
Db 181 GCAGATCCAAACATCTCCGCTCGCGCTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 240

QY 270 CTTGTCGTTATTCGGATGACGAGCGCGCTACAGCTGAGAGATGAATTCCTCTGCTGAC 329
Db 241 CTTGTCGTTATTCGGATGACGAGCGCGCTACAGCTGAGAGATGAATTCCTCTGCTGAC 300

QY 330 AACCAAAAGTGTGTTTGGAGAGGTGTTTGGAGAGGTGTTTGGAGAGGTGTTTCAAGGATC 389
Db 301 AACCAAAAGTGTGTTTGGAGAGGTGTTTGGAGAGGTGTTTGGAGAGGTGTTTCAAGGATC 360

```

```

QY 390 CACAAATGCTAAAGAGCTCTCTTCCAGGGAGTCAAAATCTGGATGCAATGGATCCG 449
Db 361 CACAAATGCTAAAGAGCTCTCTTCCAGGGAGTCAAAATCTGGATGCAATGGATCCG 420

QY 450 AGACTTTTGGACAGACCTGGGATCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAG 509
Db 421 AGACTTTTGGACAGACCTGGGATCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAG 480

QY 510 TGGCTTCAGTGGAGGAGGATTTTGGGGCAGAGATCAGAGATATGGAATCATGATTCAGG 569
Db 481 TGGCTTCAGTGGAGGAGGATTTTGGGGCAGAGATCAGAGATATGGAATCATGATTCAGG 540

QY 570 ACAGGAGATTGACCAACTCCAAAGAGTGAATGACCAATCAAAACCAACCCCTGACGACAG 629
Db 541 ACAGGAGATTGACCAACTCCAAAGAGTGAATGACCAATCAAAACCAACCCCTGACGACAG 600

QY 630 AAGAAATCATATGTGGCTTGAATCCAGAGATCTTCTCTGATGGCGCTGCCCTCCATG 689
Db 601 AAGAAATCATATGTGGCTTGAATCCAGAGATCTTCTCTGATGGCGCTGCCCTCCATG 660

QY 690 CCATCCCTCTGCGAGTCTCTATGTGTGAGACAGTGAAGTCTCTCTGATGGCGCTGCCCTCCATG 749
Db 661 CCATCCCTCTGCGAGTCTCTATGTGTGAGACAGTGAAGTCTCTCTGATGGCGCTGCCCTCCATG 720

QY 750 ATCGGAGACATGGGCTCGGTGTCCTTTCAACATCGCAGCTACGCCCTCTCTCACGTA 809
Db 721 ATCGGAGACATGGGCTCGGTGTCCTTTCAACATCGCAGCTACGCCCTCTCTCACGTA 780

QY 810 CATGATTGGCAGATCAACGGGCTGAGCCAGGTGACTTTATATACATTTGGGAGATGC 869
Db 781 CATGATTGGCAGATCAACGGGCTGAGCCAGGTGACTTTATATACATTTGGGAGATGC 840

QY 870 ACATATTTACTGTAATCAATCAGCCACTGAAATTCAGTTTTCAGTTTCAGGAGAGAGCCGACC 929
Db 841 ACATATTTACTGTAATCAATCAGCCACTGAAATTCAGTTTTCAGTTTCAGGAGAGAGCCGACC 900

QY 930 TTTCCAAAGCTCAGGATCTTTGAAAGGTTGAGA 964
Db 901 TTTCCAAAGCTCAGGATCTTTGAAAGGTTGAGA 935

RESULT 13
BQ642223 972 bp mRNA linear EST 15-JUL-2002
LOCUS BQ642223
DEFINITION BQ642223 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:6303453
5', mRNA sequence.
ACCESSION BQ642223
VERSION BQ642223.1 GI:21766395
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 972)
AUTHORS NIH-MGC http://mgi.nhl.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2522 row: i column: 22
High quality sequence stop: 590.
Location/Qualifiers
1..972
/organism="Homo sapiens"
/db_xref="taxon:9606"
FEATURES
source

```



```

DEFINITION AUI42370 Y79AA1 Homo sapiens cDNA clone Y79AA1000222 5', mRNA
sequence.
ACCESSION AUI42370
VERSION AUI42370.1 GI:11003891
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 855)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source
1..855
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Y79AA1000222"
/cell_type="retinoblastoma"
/line="v79"
/note="vector: pME18SPL3"
BASE COUNT 220 a 215 c 230 g 187 t 3 others
ORIGIN
Query Match 49.9%; Score 767; DB 9; Length 855;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 767; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 GAGCGGACCGCGCTCCGCGCAGCGGAGCTGCTAGTACTCTGGGCGAGATCAACACATC 225
DB 1 GAGCGGACCGCGCTCCGCGCAGCGGAGCTGCTAGTACTCTGGGCGAGATCAACACATC 60

QY 226 CTCGCTGGCGGCTCAGGAAGACGACCGCAGCGGACCGGACCGGACCGGACCGGACCGG 285
DB 61 CTCGCTGGCGGCTCAGGAAGACGACCGCAGCGGACCGGACCGGACCGGACCGGACCGG 120

QY 286 ATGCGGCGCGCTACAGCTCAGAGATGAATTCCTCTCTCTGACCAACAACTGTGTTC 345
DB 121 ATGCGGCGCGCTACAGCTCAGAGATGAATTCCTCTCTCTGACCAACAACTGTGTTC 180

QY 346 TGAAGAGTGTTCGAGAGTGTCTGTGTGTATCAAGGATGACCAAACTGTGTTC 405
DB 181 TGAAGAGTGTTCGAGAGTGTCTGTGTGTATCAAGGATGACCAAACTGTGTTC 240

QY 406 CTGCTCTCAAGGAGTGAATCTGGGATGCCAATGGATCCGAGACTTTTGTGGACAGC 465
DB 241 CTGCTCTCAAGGAGTGAATCTGGGATGCCAATGGATCCGAGACTTTTGTGGACAGC 300

QY 466 CTGGGATTCCTCACCAGAGAGGAGCTTGGGCGCCAGTTTATGGCTTCCAGTGAGG 525
DB 301 CTGGGATTCCTCACCAGAGAGGAGCTTGGGCGCCAGTTTATGGCTTCCAGTGAGG 360

QY 526 CATTTTGGGCGAGATACAGAGATATGAATTCAGATTTATTCAGGACAGGAGTTGACCAA 585
DB 361 CATTTTGGGCGAGATACAGAGATATGAATTCAGATTTATTCAGGACAGGAGTTGACCAA 420

QY 586 CTGCAAGAGTGAATTTGACACCATCAAAACCAACCTGACGACAGAGATCATCATGTGC 645
DB 421 CTGCAAGAGTGAATTTGACACCATCAAAACCAACCTGACGACAGAGATCATCATGTGC 480

```

```

QY 645 GCTTGGAATCCAGAGATCTTCTCTGATGGCGCTGCCATGCCATGCCCTCTTGCCAG 705
DB 481 GCTTGGAATCCAGAGATCTTCTCTGATGGCGCTGCCATGCCATGCCCTCTTGCCAG 540

QY 706 TTCTATGTGTGAACAGTGTCTCTGCCAGCTGTACAGAGATCGGGAGACATGGGC 765
DB 541 TTCTATGTGTGAACAGTGTCTCTGCCAGCTGTACAGAGATCGGGAGACATGGGC 600

QY 766 CTCGCTGTGCTTTCAACATCGCCAGCTACGCCCTGCTCAGCTACATGATGGCAGATC 825
DB 601 CTCGCTGTGCTTTCAACATCGCCAGCTACGCCCTGCTCAGCTACATGATGGCAGATC 660

QY 826 ACGGCGCTGAACAGGAGTGTCTTATACACATTTGGGAGATGCACATTTACCTGAAT 885
DB 661 ACGGCGCTGAACAGGAGTGTCTTATACACATTTGGGAGATGCACATTTACCTGAAT 720

QY 886 CACATCGAGCCACTGAAATTTACGCTTCAGCGAGAACCCAGACCTTT 932
DB 721 CACATCGAGCCACTGAAATTTACGCTTCAGCGAGAACCCAGACCTTT 767

RESULT 10
BM904778 994 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT 6699576 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557583
DEFINITION 5', mRNA sequence.
ACCESSION BM904778
VERSION BM904778.1 GI:19355157
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 994)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L14M12279 row: c column: 24
High quality sequence stop: 673.
Location/Qualifiers
1..994
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5557583"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NoI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 236 a 288 c 265 g 205 t
ORIGIN
Query Match 48.5%; Score 745; DB 14; Length 994;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 745; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 CTTTGCTTGCCTCCGTCGCCGCGCCACTTCCTGCTCCGTCCTCCGTCGCCGCGCGGC 104
DB 1 CTTTGCTTGCCTCCGTCGCCGCGCCACTTCCTGCTCCGTCCTCCGTCGCCGCGCGGC 60

QY 105 CATGCTGTGCGCGCTCGGAGCTGCGCGCGCGCCCTTCGCCCGCGCGCAGAGCG 164

```

Db	254	CAGCGCGCTACAGCCTCAGAGATGAATCCCTCTGTCTGACCAACAAACGTGTCTTCGCG	313
QY	349	AAGGGTGTCTTGGAGAGTTCCTCTGTGGTTTATCAAGGGATCCCAAAATGCTAAAGAGCTG	408
Db	314	AAGGGTGTCTTGGAGAGTTCCTCTGTGGTTTATCAAGGGATCCCAAAATGCTAAAGAGCTG	373
QY	409	TCCTTCCAAAGGGAGTGAATAATCTGGGATGCCAATGATCCCGAGACTTTTTCGACAGCCTG	468
Db	374	TCCTTCCAAAGGGAGTGAATAATCTGGGATGCCAATGAGATCTTTTTCGACAGCCTG	433
QY	469	GGATTCCTCCACAGAGAAGAAGGGGACTTGGGCCCCAGTTCCTTATGGCTTCCAGTGGAGGCAT	528
Db	434	GGATTCCTCCACAGAGAAGAAGGGGACTTGGGCCCCAGTTCCTTATGGCTTCCAGTGGAGGCAT	493
QY	529	TTTGGGCGCAATPACAGAGATATCGAATCAGATTTATCAGCACAGGAGTTGACCAACTG	588
Db	494	TTTGGGCGCAATPACAGAGATATCGAATCAGATTTATCAGCACAGGAGTTGACCAACTG	553
QY	589	CAAGAAGTGTATGACACCATCAAAACCAACCTGTGACGACAGAGAAATCATCATGTGGCT	648
Db	554	CAAGAAGTGTATGACACCATCAAAACCAACCTGTGACGACAGAGAAATCATCATGTGGCT	613
QY	649	TGGAAATCCAGAGATCTTCCTCTGATGGCGCTGCCTCCATGCCATCGCCTCTGCCAGTTC	708
Db	614	TGGAAATCCAGAGATCTTCCTCTGATGGCGCTGCCTCCATGCCATCGCCTCTGCCAGTTC	673
QY	709	TATGTGTGTGAACAGTGAGCTGTCTGCCAGCTGTACACAGAGATCGGGAGACATGGGGCTC	768
Db	674	TATGTGTGTGAACAGTGAGCTGTCTGCCAGCTGTACACAGAGATCGGGAGACATGGGGCTC	733
QY	769	GGTGTGGCTTTCAACATGCCAGCTACGCCCTGCTCAGTACATGATTCGCGACATCAACG	828
Db	734	GGTGTGGCTTTCAACATGCCAGCTACGCCCTGCTCAGTACATGATTCGCGACATCAACG	793
QY	829	GGCCTGAAGCCAGG	842
Db	794	GGCCTGAAGCCAGG	807

```

RESULT 8
LOCUS BM473957
DEFINITION AGENCOURT_6467897 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:582212
5', mRNA sequence.
ACCESSION BM473957
VERSION BM473957.1 GI:18522999
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1035)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12343 row: f column: 05
High quality sequence stop: 687.
FEATURES
source
location/Qualifiers
1..1035
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:582212"
/cloae lib="NHT MGC 67"

```

/tissue type="placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 220 a 263 c 254 g 193 t 2 others
ORIGIN

Query Match 52.0%; Score 799; DB 9; Length 932;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 899; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 66 CAGGCACTTCCCTGCTCGTCCGCCCCCGCGCGCCAGTCCGCGGCTCGGA 125
Db 1 CCGCGCACTTCCCTGCTCGTCCGCCCCCGCGCGCCAGTCCGCGGCTCGGA 60
QY 126 GCTGCGCGCGCGCCCTTGGCCCCCGCGCGCCAGTCCGCGGCTCGGCC 185
Db 61 GCTGCGCGCGCGCCCTTGGCCCCCGCGCGCCAGTCCGCGGCTCGGCC 120
QY 186 GCACGGGAGCTGCAGTACCTTGGGGCAGATCCACATCTCCGCTGGCGGTACAGAA 245
Db 121 GCACGGGAGCTGCAGTACCTTGGGGCAGATCCACATCTCCGCTGGCGGTACAGAA 180
QY 246 GGAACGCGACCGGCGACCGGCGACCTGTCGCTATTCGGATGACGCGCGCTACAGCT 305
Db 181 GGAACGCGACCGGCGACCGGCGACCTGTCGCTATTCGGATGACGCGCGCTACAGCT 240
QY 306 GAGATGAATTCCTCTGCTGACCAACCAAGCTGTGTCTGGAAGGCTTTTGGAGGA 365
Db 241 GAGATGAATTCCTCTGCTGACCAACCAAGCTGTGTCTGGAAGGCTTTTGGAGGA 300
QY 366 GTTGTGTGTTTATCAAGGATCCAAATGCTAAAGAGCTGTCTTCCAAAGGAGTGAA 425
Db 301 GTTGTGTGTTTATCAAGGATCCAAATGCTAAAGAGCTGTCTTCCAAAGGAGTGAA 360
QY 426 AATCTGGGATCCCAATGATCCCGAGACTTTTGGACAGCTTGGGATTCCTCCACAGAGA 485
Db 361 AATCTGGGATCCCAATGATCCCGAGACTTTTGGACAGCTTGGGATTCCTCCACAGAGA 420
QY 486 AGAAGGGAGCTTGGGCCCGAGTTATGGCTTCCAGTGAGGAGATTTGGGGCAGAAATACAG 545
Db 421 AGAAGGGAGCTTGGGCCCGAGTTATGGCTTCCAGTGAGGAGATTTGGGGCAGAAATACAG 480
QY 546 AGATATGAATCAGATTTTACAGACAGGAGTTGACCACTGCAAGAGTGTATTCACAC 605
Db 481 AGATATGAATCAGATTTTACAGACAGGAGTTGACCACTGCAAGAGTGTATTCACAC 540
QY 606 CATCAAAACCAACCTGACACAGAGATTCATGTCGCTTGGATCCAGAGATCT 665
Db 541 CATCAAAACCAACCTGACACAGAGATTCATGTCGCTTGGATCCAGAGATCT 600
QY 666 TCCTGTGATGGCGTGTCTCCATGCCATGCCCTCTGCCAGTTCTATGTGTGAAACAGTGA 725
Db 601 TCCTGTGATGGCGTGTCTCCATGCCATGCCCTCTGCCAGTTCTATGTGTGAAACAGTGA 660
QY 726 GCTGTCTCCAGCTGTACACAGATCGGAGACATGGGCGCTTCGCTGTGCTTTCAACAT 785
Db 661 GCTGTCTCCAGCTGTACACAGATCGGAGACATGGGCGCTTCGCTGTGCTTTCAACAT 720
QY 786 CGCAGCTACGCGCTGTCTACATCATGATTCGACATTCACGCGGCTTGAAGCAGGTGA 845
Db 721 CGCAGCTACGCGCTGTCTACATCATGATTCGACATTCACGCGGCTTGAAGCAGGTGA 780
QY 846 CTTTATACACATTTGGGAGATGCATATTTTACCTGAATCATTCGAGGCACTGAAAT 905
Db 781 CTTTATACACATTTGGGAGATGCATATTTTACCTGAATCATTCGAGGCACTGAAAT 840

QY 906 TCAGTTTCAGCGAGACCCAGACCTTTCCCAAGCTCAGGATTTCTTGGAAAAGTTGAGAA 965
Db 841 TCAGTTTCAGCGAGACCCAGACCTTTCCCAAGCTCAGGATTTCTTGGAAAAGTTGAGAA 900
QY 966 A 966
Db 901 A 901

RESULT 7
BQ420981
LOCUS BQ420981 889 bp mRNA linear EST 23-MAY-2002
DEFINITION AGSCOURT_7761945 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6018140
5', mRNA sequence.
ACCESSION BQ420981
VERSION BQ420981.1 GI:21116296
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 889)
NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLMI3218 row: i column: 21
High quality sequence stop: 715.

FEATURES
source
1..889
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:6018140"
/clone_lib="NIH MGC 70"
/tissue_type="epithelioid carcinoma"
/lab host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life Technologies."
BASE COUNT 195 a 265 c 248 g 181 t
ORIGIN

Query Match 51.7%; Score 794; DB 14; Length 889;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 794; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 GCCTGTGCTCGTCCGCGCGCACCTTCGCTCCCTCCGCGCGCGCGCGCGCATG 108
Db 14 GCCTGTGCTCGTCCGCGCGCACCTTCGCTCCCTCCGCGCGCGCGCGCGCATG 73
QY 109 CCTGTGCGCGGTCTCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGGAC 168
Db 74 CCTGTGCGCGGTCTCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGGAC 133
QY 169 GCCGAGCGCGGTCCGCGCGCACGGGAGCTGACGACTCTGGGCGACATCCACATCTC 228
Db 134 GCCGAGCGCGGTCCGCGCGCACGGGAGCTGACGACTCTGGGCGACATCCACATCTC 193
QY 229 CGCTGCGCGGTTCAGGAAGAGCAGCGACCGGGCACCAGCCCTGTGCGTATTCGCGATG 288
Db 194 CGCTGCGCGGTTCAGGAAGAGCAGCGACCGGGCACCAGCCCTGTGCGTATTCGCGATG 253
QY 289 CAGCGCGCGCTACAGCGCTGAGAGATGAATTCCTCTGCTGACCAACCAACGTTCTGG 348

```

Db 661 CTTGCGAGCTGTACAGAGATCGGAGACATGGGCTCGGTGCTTTCAACATCGCA 720
Qy 791 GCTACGCCCTGCTACAGTACATGATGCGCATACACGGGCTGAGACCGAGTACTTTA 850
Db 721 GCTACGCCCTGCTACAGTACATGATGCGCATACACGGGCTGAGACCGAGTACTTTA 780
Qy 851 TACACACTTTGGGAGATGCAATATTTACCTGATCAATCATGAGCCACTGAAATTCAGC 910
Db 781 TACACACTTTGGGAGATGCAATATTTACCTGATCAATCATGAGCCACTGAAATTCAGC 840
Qy 911 TTTACGCGAGAACCC 924
Db 841 TTTACGCGAGAACCC 854

RESULT 5
BM478045 1045 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6482920 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5552680
5', mRNA sequence.
ACCESSION BM478045.1 GI:18527087
VERSION BM478045.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1045)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12268 row: 9 column: 17
High quality sequence stop: 687.
FEATURES
Location/Qualifiers
1..1045
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5552680"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="organ: lymph, Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."
BASE COUNT 242 a 297 c 286 g 220 t
ORIGIN
Query Match 53.7%; Score 825; DB 13; Length 1045;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 825; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 33 CCCGCGCGGCACATGGCTGCTCCGTCGCGCGCGCCACCTTCGCTGCTCGTCCCC 92
Db 1 CCGCGCGCGGCACATGGCTGCTCCGTCGCGCGCGCCACCTTCGCTGCTCGTCCCC 60
Qy 93 CGCCCGCGCGGCACATGGCTGCTGCGCGCGGTGCGAGCTGCGCGCGCGCCCTTCCCGCCCGC 152
Db 61 CGCCCGCGCGGCACATGGCTGCTGCGCGCGGTGCGAGCTGCGCGCGCGCCCTTCCCGCCCGC 120
Qy 153 CGCAGAGAGCGGAGCGCGAGCGCGGTGCTCCGCGCGAGCGGAGACTCAGTACTGGGCA 212

```

```

Db 121 CGCACAGGAGCGGAGCGCGAGCCGCTCCCGCCGACGCGGAGAGTGCAGTACCTGGGGCA 180
Qy 213 GATCCAAACATCTCTCCGCTCGGCGGTGAGAAAGACGACCGCAGCGGCGACCGGACACCT 272
Db 181 GATCCAAACATCTCTCCGCTCGGCGGTGAGAAAGACGACCGCAGCGGCGACCGGACACCT 240
Qy 273 GTCGGTATTCGGCATGCGAGCGCGCTACAGCTGAGAGATGAATTCCTCTGCTGACAAAC 332
Db 241 GTCGGTATTCGGCATGCGAGCGCGCTACAGCTGAGAGATGAATTCCTCTGCTGACAAAC 300
Qy 333 CAAACGCTGCTCTGGAAGGGTGTGTTGGAGGAGTGTGCTGTGTTATCAAGGGATCCAC 392
Db 301 CAAACGCTGCTCTGGAAGGGTGTGTTGGAGGAGTGTGCTGTGTTATCAAGGGATCCAC 360
Qy 393 AAATGCTAAAGAGCTGTCTTCCAGGGAGTGAAATCTGGATGCCAATGGATCCCGAGA 452
Db 361 AAATGCTAAAGAGCTGTCTTCCAGGGAGTGAAATCTGGATGCCAATGGATCCCGAGA 420
Qy 453 CTTTTCGACAGCGCTGGGATTTCTCCACAGAGAGAGAGGGGACTTGGGCCCGAGTTTATGG 512
Db 421 CTTTTCGACAGCGCTGGGATTTCTCCACAGAGAGAGAGGGGACTTGGGCCCGAGTTTATGG 480
Qy 513 CTTCCAGTGGAGGCAATTTTGGGGCAGATATCAGAGATATGGAATCAGATATTACGAGACA 572
Db 481 CTTCCAGTGGAGGCAATTTTGGGGCAGATATCAGAGATATGGAATCAGATATTACGAGACA 540
Qy 573 GGGAGTTGACCAACTGCAAGAGTGAATTCACCATCAAAACCAACCTTCGACGACAGAG 632
Db 541 GGGAGTTGACCAACTGCAAGAGTGAATTCACCATCAAAACCAACCTTCGACGACAGAG 600
Qy 633 AATCATCATGTGCGCTTGAATTCGAAGAGATCTTCCTGATGAGGCGCTGCTCCATGCCA 692
Db 601 AATCATCATGTGCGCTTGAATTCGAAGAGATCTTCCTGATGAGGCGCTGCTCCATGCCA 660
Qy 693 TGGCCTCTGCCAGTTCTATGTGTGTGAACAGTGAAGTGTCTCTGAGTGGCTGCTGACGAGATC 752
Db 661 TGGCCTCTGCCAGTTCTATGTGTGTGAACAGTGAAGTGTCTCTGAGTGGCTGCTGACGAGATC 720
Qy 753 GGGAGATATGGCGCTGGGTGCTTCACATCGCAGCTGCTGCGAGCTGCTGACGCTGACGAGATC 812
Db 721 GGGAGATATGGCGCTGGGTGCTTCACATCGCAGCTGCTGCGAGCTGCTGACGCTGACGATAT 780
Qy 813 GATTGCGCACATCAACGGGCTGAGACCGAGTGAATTTATACACAC 857
Db 781 GATTGCGCACATCAACGGGCTGAGACCGAGTGAATTTATACACAC 825

RESULT 6
AL551817 932 bp mRNA linear EST 16-FEB-2001
LOCUS AL551817 LFI_NFL006_PL2 Homo sapiens cDNA clone CS0D1061YF09 5
DEFINITION prime, mRNA sequence.
ACCESSION AL551817
VERSION AL551817.1 GI:12890127
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 932)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..932
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1061YF09"
/clone_lib="LFI_NFL006_PL2"
FEATURES
source

```


QY 74 TTGCGCTGCTCCGTCCTCCCGCCCGCCGCGCATGCTGTGCTCCGCTCGAGCTCGCG 133
 Db 67 TTGCGCTGCTCCGTCCTCCCGCCCGCCGCGCATGCTGTGCTCCGCTCGAGCTCGCG 126
 QY 134 GCCGGCCCTTGGCCCTCCCGCCCGCCGCGCGAGCGGAGCGCCGCTCGCGCGAGGG 193
 Db 127 GCCGGCCCTTGGCCCTCCCGCCCGCCGCGCGAGCGGAGCGCCGCTCGCGCGAGGG 186
 QY 194 AGCTGAGTACCTGGGCGAGTCCACACATCTCTCGCTGGGCGCTCAGGAGGAGGACC 253
 Db 187 AGCTGAGTACCTGGGCGAGTCCACACATCTCTCGCTGGGCGCTCAGGAGGAGGACC 246
 QY 254 GCACGGGACCGGACCTGCTGCGGTATTTGGCATGAGCGCGCGCTACAGCTCAGAGATG 313
 Db 247 GCACGGGACCGGACCTGCTGCGGTATTTGGCATGAGCGCGCGCTACAGCTCAGAGATG 306
 QY 314 AATTCCTCTGCTGACACCAACAAAGTGTGTTCTGGAAGGGTGTGTTGGAGGAGTGTCT 373
 Db 307 AATTCCTCTGCTGACACCAACAAAGTGTGTTCTGGAAGGGTGTGTTGGAGGAGTGTCT 366
 QY 374 GGTATCAGGAGTCCACAAATGCTAAAGAGTGTCTTCAAGGAGTGAATACTGGG 433
 Db 367 GGTATCAGGAGTCCACAAATGCTAAAGAGTGTCTTCAAGGAGTGAATACTGGG 426
 QY 434 ATGCCAATGGATCCGAGACTTTTTCGACAGCTCGGATCTCCACGAGAGAGAGGG 493
 Db 427 ATGCCAATGGATCCGAGACTTTTTCGACAGCTCGGATCTCCACGAGAGAGAGGG 486
 QY 494 ACTTGGGCGCCAGTTTATGGCTTCAGTGGAGGCAATTTTGGGCGAGATACAGAGATATGG 553
 Db 487 ACTTGGGCGCCAGTTTATGGCTTCAGTGGAGGCAATTTTGGGCGAGATACAGAGATATGG 546
 QY 554 AATCAGATATTTCAGGACAGGAGTTGACCAACTGCAAGAGTGAATTGACACCTCAAAA 613
 Db 547 AATCAGATATTTCAGGACAGGAGTTGACCAACTGCAAGAGTGAATTGACACCTCAAAA 606
 QY 614 CCAACCTGAGCAGAGAGATCATCATGCTGGCTTGGAAATCCAGAGATCTTCTCTGTA 673
 Db 607 CCAACCTGAGCAGAGAGATCATCATGCTGGCTTGGAAATCCAGAGATCTTCTCTGTA 666
 QY 674 TGGCGCTGCTCCATGCGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 733
 Db 667 TGGCGCTGCTCCATGCGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 726
 QY 734 GCCAGCTTACAGAGATCGGAGACATGGGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
 Db 727 GCCAGCTTACAGAGATCGGAGACATGGGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 786
 QY 794 AGCCCTGCTCAGTACATGATGCGACATCATCGGCTTGAAGCCAGGCTGATTTATAC 853
 Db 787 AGCCCTGCTCAGTACATGATGCGACATCATCGGCTTGAAGCCAGGCTGATTTATAC 846
 QY 854 ACACCTTTGGGAGATGACATATTTACTGAATCATCGAGGCGACTGA 901
 Db 847 ACACCTTTGGGAGATGACATATTTACTGAATCATCGAGGCGACTGA 894

RESULT 4
 AL556369
 LOCUS
 DEFINITION AL556369 LTI_NFL006 PL2 Homo sapiens cDNA clone CS0DK004YL08 5
 prime, mRNA sequence.
 ACCESSION AL556369
 VERSION AL556369.1 GI:12898981
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 854)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)

COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 1. .854
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DK004YL08"
 /clone_lib="LTI_NFL006_PL2"
 /issue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact : Feng Liang Life Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"
 BASE COUNT 194 a 247 c 239 g 174 t
 ORIGIN
 Query Match 55.6%; Score 854; DB 9; Length 854;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 854; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 71 CACTTCGCTGCTCCGTCCTCCCGCCCGCCGCGCATGCTGTGCGCGCTCGAGCTGC 130
 Db 1 CACTTCGCTGCTCCGTCCTCCCGCCCGCCGCGCATGCTGTGCGCGCTCGAGCTGC 60
 QY 131 CGCGCGCGCTTGGCCCTCCCGCCCGCCGCGAGCGGAGCGCCGCGCTCCCGCCGACG 190
 Db 61 CGCGCGCGCTTGGCCCTCCCGCCCGCCGAGCGGAGCGCGCGCTCCCGCCGACG 120
 QY 191 GGGAGCTGCTAGTACCTGGGCGGAGATCCAAACATCTCTCGCTCGGCGCTCAGGAGGACG 250
 Db 121 GGGAGCTGCTAGTACCTGGGCGGAGATCCAAACATCTCTCGCTCGGCGCTCAGGAGGACG 180
 QY 251 ACCGACGGGCGACCGGCACTGTCGCTATTTGGGCGATGACGCGCGCTTACAGCTCAGAG 310
 Db 181 ACCGACGGGCGACCGGCACTGTCGCTATTTGGGCGATGACGCGCGCTTACAGCTCAGAG 240
 QY 311 ATGAATTCCTCTGCTGACCAACCAACGCTGTCTTCTGGAAGGGTGTGTTGGAGGAGTGC 370
 Db 241 ATGAATTCCTCTGCTGACCAACCAACGCTGTCTTCTGGAAGGGTGTGTTGGAGGAGTGC 300
 QY 371 TGTGGTTTATCAAGGATCCAAATGCTAAAGAGCTGTCTTCCAGGAGTGAATCT 430
 Db 301 TGTGGTTTATCAAGGATCCAAATGCTAAAGAGCTGTCTTCCAGGAGTGAATCT 360
 QY 431 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGCTTCCACGAGAGAGAG 490
 Db 361 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGCTTCCACGAGAGAGAG 420
 QY 491 GGGACTTGGGCGCCAGTTTATGCTTCCAGTGGAGGCAATTTTGGGCGAGAAATACAGAGATA 550
 Db 421 GGGACTTGGGCGCCAGTTTATGCTTCCAGTGGAGGCAATTTTGGGCGAGAAATACAGAGATA 480
 QY 551 TGGATCAGATATTTCAGGACAGGAGTGAACCAATGCAAGAGTGAATGACACCATCA 610
 Db 481 TGGATCAGATATTTCAGGACAGGAGTGAACCAATGCAAGAGTGAATGACACCATCA 540
 QY 611 AAACCAACCTGCGACGACAGAGATCATGTCGCTTGGATCCAGAGATCTTCTC 670
 Db 541 AAACCAACCTGCGACGACAGAGATCATGTCGCTTGGATCCAGAGATCTTCTC 600
 QY 671 TGATGGCGCTGCTCCATCCCTGCTCCAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCT 730
 Db 601 TGATGGCGCTGCTCCATCCCTGCTCCAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 QY 731 CCGCCAGCTGTACAGAGATCGGAGAGATGCGGCTCGGCTGCTGCTGCTGCTGCTGCTGCT 790

JOURNAL
COMMENT

Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers

FEATURES
source

1. .1020
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI008YH21"
/clone_lib="LRI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 237 a 291 c 276 g 215 t 1 others
ORIGIN

Query Match 64.1%; Score 985; DB 9; Length 1020;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 GCCTGCTCGTCCGCGCGCCACTGCGCTGCTCGTCCGCGCGCGCGCGCCATG 108
Db 15 GCCTGCTCGTCCGCGCGCCACTGCGCTGCTCGTCCGCGCGCGCGCGCCATG 74
QY 109 CTTGTCGCGCTCGAGAGTGCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGAC 168
Db 75 CTTGTCGCGCTCGAGAGTGCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGAC 134
QY 169 GCGGAGCGCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 228
Db 135 GCGGAGCGCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 194
QY 229 CGCTGCGCGCTCAGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 288
Db 195 CGCTGCGCGCTCAGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 254
QY 289 CAGGCGCGCTACGCTGAGAGTGAATTCCTGCTGCAACCAACCAACCAACG 348
Db 255 CAGGCGCGCTACGCTGAGAGTGAATTCCTGCTGCAACCAACCAACCAACG 314
QY 349 AAGGCGTGTGTTGAGAGGAGTGTGTTGTTATCAAGGAGTCCCAAAATGCTAA 408
Db 315 AAGGCGTGTGTTGAGAGGAGTGTGTTGTTATCAAGGAGTCCCAAAATGCTAA 374
QY 409 TCTTCAAGGAGTGAATAATCTGGAGTGCATGATCCGAGACTTTTGGAGAGCTG 468
Db 375 TCTTCAAGGAGTGAATAATCTGGAGTGCATGATCCGAGACTTTTGGAGAGCTG 434
QY 469 GGAATTCACACAGAGAGGAGGAGTGGCCCGAGTTATGCTTCAGTGGAGGAT 528
Db 435 GGAATTCACACAGAGAGGAGGAGTGGCCCGAGTTATGCTTCAGTGGAGGAT 494
QY 529 TTTGGGCGAGATACAGAGATATGGAATCAGATTATTCAGACAGAGGAGTGA 588
Db 495 TTTGGGCGAGATACAGAGATATGGAATCAGATTATTCAGACAGAGGAGTGA 554
QY 589 CAAAGAGTATTGACACATCAAAACCAACCTGACAGACAGAGAAATCATGTCGCT 648
Db 555 CAAAGAGTATTGACACATCAAAACCAACCTGACAGACAGAGAAATCATGTCGCT 614
QY 649 TGAATTCAGAGAGTCTTCTCTGATGGCGCTGCTCCATGCCATGCCCTTGCACG 708
Db 615 TGAATTCAGAGAGTCTTCTCTGATGGCGCTGCTCCATGCCATGCCCTTGCACG 674

QY 709 TATGTGTGAACAGTGAAGCTGCTCTCGAGCTGTACACAGATCGGAGACATGGCCCTC 768
Db 675 TATGTGTGAACAGTGAAGCTGCTCTCGAGCTGTACACAGATCGGAGACATGGCCCTC 734
QY 769 GGTGTGCTTTTCAACATCGCCAGCTAGCCCTGCTCAGTACATGATTGGCACATCA 828
Db 735 GGTGTGCTTTTCAACATCGCCAGCTAGCCCTGCTCAGTACATGATTGGCACATCA 794
QY 829 GGCCTGAAGCCAGCTGACTTTATACACACTTTGGGAGATGCACATATTACCTGAATCAC 888
Db 795 GGCCTGAAGCCAGCTGACTTTATACACACTTTGGGAGATGCACATATTACCTGAATCAC 854
QY 889 ATCGAGCCACTGAAATTTAGCTTTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGATT 948
Db 855 ATCGAGCCACTGAAATTTAGCTTTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGATT 914
QY 949 CTTGAAAGTTGAGAAATTTAGTACTTTCAAAGCTGAAGACTTTTCAGATTGAAGGTTAC 1008
Db 915 CTTGAAAGTTGAGAAATTTAGTACTTTCAAAGCTGAAGACTTTTCAGATTGAAGGTTAC 974
QY 1009 AATCGCATCCACTATTAATGG 1033
Db 975 AATCGCATCCACTATTAATGG 999

RESULT 3
AL551990 938 bp mRNA linear EST 16-FEB-2001
LOCUS AL551990 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI060YE04 5
DEFINITION prime, mRNA sequence.

ACCESSION AL551990
VERSION AL551990.1 GI:12890459
KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 938)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source

1. .938
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI060YE04"
/clone_lib="LRI_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 202 a 288 c 259 g 188 t 1 others
ORIGIN

Query Match 57.8%; Score 888; DB 9; Length 938;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 888; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 CCACTTGGCTGCTCGTCCGCGCGCGCGCGCGCGCTGGCTCGCTCGCGCGCGCGCAC 73
Db 7 CCACTTGGCTGCTCGTCCGCGCGCGCGCGCGCGCTGGCTCGCTCGCGCGCGCAC 66

Shevchenko, Y., Metherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Jagsapi, R., Lim, M., Madhro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantrop, S., Thomas, P.J., Thompson, E.B., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/INL at: <http://image.lnl.gov>
Series: IRAL plate: 4 Row: h Column: 10
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 10635001
This clone has the following problem: no 5' EST match.

FEATURES

source

1.1533

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3141419"

/tissue_type="Ovary, adenocarcinoma"

/clone_lib="NIH MGC 9"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

BASE COUNT

424 a 354 c 382 g 373 t

ORIGIN

Query Match

84.0%; Score 1291; DB 11; Length 1533;

Best Local Similarity

99.8%; Pred. No. 0;

Matches 1441; Conservative

0; Mismatches 3; Indels 0; Gaps 0;

53 GCTTCCTGCTCCGCGCGGCACTTTCGCTCCGCTCCGCGCGCGCGCGCGCATGCTG 112
 9 GCTTCCTGCTCCGCGCGGCACTTTCGCTCCGCTCCGCGCGCGCGCGCGCATGCTG 68
 113 TGCCCGGCTGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATGCTG 172
 69 TGCCCGGCTGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATGCTG 128
 173 AGCCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATGCTGCT 232
 129 AGCCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATGCTGCT 188
 233 GCGCGCTGCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATGCTG 292
 189 GCGCGCTGCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATGCTG 248
 293 GCGCGCTGCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATGCTG 352
 249 GCGCGCTGCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATGCTG 308
 353 GTGTTTGGAGAGGAGTGTGTTTATCAAGGATCCAAATGCTAAAGAGCTGCTT 412
 309 GTGTTTGGAGAGGAGTGTGTTTATCAAGGATCCAAATGCTAAAGAGCTGCTT 368
 413 CCAGAGGAGTGAATCTGGAATGCCAATGATCCCAAGCTTTTGGACAGCTGGAGT 472
 369 CCAGAGGAGTGAATCTGGAATGCCAATGATCCCAAGCTTTTGGACAGCTGGAGT 428
 473 TCTCCACCAAGAAAGAGGAGCTTGGGCGCACTTATGCTTCAAGGAGCACTTTTG 532
 429 TCTCCACCAAGAAAGAGGAGCTTGGGCGCACTTATGCTTCAAGGAGCACTTTTG 488
 533 GGGCGAATACAGATATGGAATCAATATCAAGGAGGAGGAGGAGGAGGAGGAGGAG 592
 489 GGGCGAATACAGATATGGAATCAATATCAAGGAGGAGGAGGAGGAGGAGGAGGAG 548
 593 GAGGATTTGACACATCAAAACCAACCTGAGCGACAGAGAGATCAATCAATGAGCTG 652
 549 GAGGATTTGACACATCAAAACCAACCTGAGCGACAGAGAGATCAATCAATGAGCTG 608
 653 ATCCAGAGATCTTCTGATGGCGCTGCTCCATGCGAGGAGGAGGAGGAGGAGGAGGAG 712
 609 ATCCAGAGATCTTCTGATGGCGCTGCTCCATGCGAGGAGGAGGAGGAGGAGGAGGAG 668

713 TGTGAACAGTGAAGTCTGCTGCGAGCTGTAACAGAGATCGGAGACATGCGCTCGGTG 772
 669 TGTGAACAGTGAAGTCTGCTGCGAGCTGTAACAGAGATCGGAGACATGCGCTCGGTG 728
 773 TGCCTTCAACATCGCAGCTACGCGCTGCTGCTACAGTACATGATTCGGACATCAGCGGCC 832
 729 TGCCTTCAACATCGCAGCTACGCGCTGCTGCTACAGTACATGATTCGGACATCAGCGGCC 788
 833 TGAAGCAGGTGACTTATACACACTTGGAGATGACATATTTACAGTAACAGATCG 892
 789 TGAAGCAGGTGACTTATACACACTTGGAGATGACATATTTACAGTAACAGATCG 848
 893 AGCCATGGAATTTGAGCTTGAAGAGAACCCAGACCTTTCCCAAGCTCAGATCTTC 952
 849 AGCCATGGAATTTGAGCTTGAAGAGAACCCAGACCTTTCCCAAGCTCAGATCTTC 908
 953 GAAAGTTGAGAAATTTGATGATGATCAAGCTGAGACCTTCAATGAGAGGATCAATC 1012
 909 GAAAGTTGAGAAATTTGATGATGATCAAGCTGAGACCTTCAATGAGAGGATCAATC 968
 1013 CGCATCCACTATTAATAATGAAATGCTGTTTGGTGTCTTTCAAGAGCTGAGAGGA 1072
 969 CGCATCCACTATTAATAATGAAATGCTGTTTGGTGTCTTTCAAGAGCTGAGAGGA 1028
 1073 TATGTGAGTCTTTAGGAGGTGGGCTGATGCGAGAGTAAAGTCTTTGCTCTTAAA 1132
 1029 TATGTGAGTCTTTAGGAGGTGGGCTGATGCGAGAGTAAAGTCTTTGCTCTTAAA 1088
 1133 GAAAGAGAACTGATGCTAAATTCGCTGACCTTCAATGATTAATTTTAAGATG 1192
 1089 GAAAGAGAACTGATGCTAAATTCGCTGACCTTCAATGATTAATTTTAAGATG 1148
 1193 TTGCACTGCGAAATGATGATGCTGCAATCTTTCATTAATAAGCTTTGATTAATCT 1252
 1149 TTGCACTGCGAAATGATGATGCTGCAATCTTTCATTAATAAGCTTTGATTAATCT 1208
 1253 CACTGAGGATATGCAATGCTGAGGTTTGAACAAAGTGAAGAGATGATGATG 1312
 1209 CACTGAGGATATGCAATGCTGAGGTTTGAACAAAGTGAAGAGATGATGATG 1268
 1313 GCTTTAGCAAAAATCATGATGCTGATTTCAATCCACGATCTTAATAAGAGTTGGTG 1372
 1269 GCTTTAGCAAAAATCATGATGCTGATTTCAATCCACGATCTTAATAAGAGTTGGTG 1328
 1373 AATTTCACAGCTATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 1432
 1329 AATTTCACAGCTATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 1388
 1433 CCTCAATTTGAGGAGCTGATGATCAACATGATGATGATGATGATGATGATGATGATG 1492
 1389 CCTCAATTTGAGGAGCTGATGATCAACATGATGATGATGATGATGATGATGATGATG 1448
 1493 TTTA 1496
 1449 TTTA 1452

RESULT 2
 LOCUS AL547612 1020 bp mRNA linear EST 16-FEB-2001
 DEFINITION AL547612 LTI NF1006_P12 Homo sapiens cDNA clone CS010087H21 5
 prime, mRNA sequence.
 ACCESSION AL547612
 VERSION AL547612.1 GI:12881832
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1020)
 AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Playes, D.
 TITLE Full-length cDNA libraries and normalization

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 29, 2002, 00:16:50 ; Search time 2021.68 Seconds
(without alignments)
12304.774 Million cell updates/sec

Title: US-09-963-333-6
Perfect score: 1536
Sequence: 1 ggggggggggggggaccacttg.....ataataagaagtgtctgc 1536

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 15

Total number of hits satisfying chosen parameters: 36158

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_estl:*
10: gb_est2:*
11: gb_estc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1291	84.0	1533	11	BC018858 Homo sapi
2	985	64.1	1020	9	AL547612
3	888	57.8	938	9	AL551990
4	854	55.6	854	9	AL556369
5	825	53.7	1045	13	BM478045
6	799	52.0	932	9	AL551817

7	794	51.7	889	14	BQ420981
8	781	50.8	1035	13	BM473957
9	767	49.9	855	9	AU142370
10	745	48.5	994	14	BM904778
11	736	47.9	948	14	BQ068205
12	731	47.6	935	9	AL542409
13	712	46.4	972	14	BQ642223
14	705	45.9	888	12	BG575942
15	696	45.3	1071	14	BQ057866
16	695	45.2	903	13	BM459789
17	690	44.9	1089	9	AL575295
18	688	44.8	1033	14	BM928567
19	687	44.7	742	9	AL548721
20	682	44.4	722	12	BG574546
21	682	44.4	900	13	BM453660
22	675	43.9	1046	14	BM919886
23	669	43.6	1072	13	BM475928
24	657	42.8	809	9	AU117646
25	656	42.7	1072	14	BQ066271
26	653	42.5	904	14	BQ642214
27	651	42.4	841	14	BQ434233
28	650	42.3	1149	13	BM546045
29	646	42.1	907	14	BQ223405
30	643	41.9	888	13	BM413631
31	641	41.7	939	13	BM452851
32	638	41.5	1135	14	BQ062012
33	634	41.3	893	13	BM457044
34	630	41.0	987	13	BM472540
35	629	41.0	789	13	BI871083
36	623	40.6	904	14	BQ669537
37	623	40.6	970	14	BQ928426
38	617	40.2	741	9	AU118644
39	614	40.0	984	14	BQ061952
40	613	39.9	890	12	BG107998
41	609	39.6	891	14	BQ224163
42	602	39.2	825	9	AU117269
43	602	39.2	906	9	AL578646
44	595	38.7	738	9	AU130205
45	591	38.5	591	14	BM843120

ALIGNMENTS

BC018858 Homo sapiens, clone IMAGE:3141419, mRNA. 1533 bp linear HTC 07-DEC-2001
LOCUS BC018858
DEFINITION BC018858
ACCESSION BC018858
VERSION BC018858.1 GI:17402996
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (Bases 1 to 1533)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: DCTD/DP
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov

QY 361 GAGAGTTGCTGTGGTTTATCAAGGATCCACAAATGCTAAAGAGCTGTCTTCCAGGGA 420
Db 404 GAGAGTTGCTGTGGTTTATCAAGGATCCACAAATGCTAAAGAGCTGTCTTCCAGGGA 463
QY 421 GTGAAATCTGGGATGCCAATCGATCCCGAGCTTTTGGACAGCTGGGATTTCCACC 480
Db 464 GTGAAATCTGGGATGCCAATCGATCCCGAGCTTTTGGACAGCTGGGATTTCCACC 523
QY 481 AGAAGAGAGGGGACTGGGCCAGTTTATGGCTTCCAGTGGAGGCAATTTGGGGCAGAA 540
Db 524 CGACAGAGAGGGGACTGGGCCAGTTTATGGCTTCCAGTGGAGGCAATTTGGGGCAGAG 583
QY 541 TACAGAGATGGAATCAGATTAATTCAGGACAGGAGTTGACCAACTGCCAAAGAGTGAAT 600
Db 584 TACAAAGACATGGAATCAGATTAATTCAGGACAGGAGTTGACCAACTGCCAAAGAGTGAAT 643
QY 601 GACACCATCAAAACCAACCTGACACAGAGATCATCATGTGGCTTGGAAATCCAGA 660
Db 644 GACACCATCAAAACCAACCTGATGACAGAGATCATCATGTGGCTTGGAAATCCAAA 703
QY 661 GATCTTCTCTGATGGG---CTGCTCCATGCCATGCCCTCTGCCAGTTCTATGTGT 716
Db 704 GATCTTCTCTGATGGGACTGGCTGCTCTCTGCCATGCCCTCTGTCACTTCTATGTGT 763
QY 717 GACAGTGTGCTGTCTGCGAGCTGTACAGAGATCGGAGACATGGGCTCGGTGTGCC 776
Db 764 GAATGGGGAGCTATCTGTGCGAGCTTTACAGAGGTGAGGAGATGAGGTCTGGGTGTGCC 823
QY 777 TTTCACATCGCAGCTAGCCCTGCTCAGCTACATGATTCGCAATCACGGGCTG-- 834
Db 824 TTTCACATCGCAGCTATGCTCTGCTCAGCTACATGATTCGCAATCACGGGCTGCA 883
QY 835 -----AAGCCAGGTGACTTTATACACACTTTGGGAGATGCACATATTTACCTGAATC 886
Db 884 GACATTTTAGCCAGGTGATTTTGTCCACAAATTTGGGAGATGCACATATTTACCTGAATC 943
QY 887 ACATGAGCCACTGAAATTCAGCTTCAGCGAGAACCCAGACCTTTCCCAAAGCTCAGGA 946
Db 944 ATATAGAGCCGCTGAAATTCAGCTACAGG-----AGACCTTTGCCAAAGCTCAAAA 996
QY 947 TTCTTGGAAAGTTGAGAAATTCAGCTTCAGGAGCTTTCAAGCTTTTCAGATTTGAGGGT 1006
Db 997 TCCTTGGAAAGTTGAGCAATTCAGCTTCAGGAGCTTTCAAGCTTTTCAGATTTGAGGGT 1056
QY 1007 ACAATCCGCTCCAACTTAATTAAGGAAATGCTGTTTGGGTGCTTTCAAAGGAGCTN 1066
Db 1057 ATAATCCACATCCCAACGATTAAGGAGATGCTGTTTGGAGAGCTTTTCAGTGTGCTG 1116
QY 1067 GAAGGAT 1073
Db 1117 TGAATAT 1123

Search completed: November 28, 2002, 21:40:03
Job time : 3915.28 secs

SOURCE	Rattus norvegicus.	
ORGANISM	Rattus norvegicus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
REFERENCE	1 (bases 1 to 1598)	
AUTHORS	Ciesla,J., Weiner,K.X., Weiner,R.S., Reston,J.T., Maley,G.F. and Maley,F.	
TITLE	Isolation and expression of rat thymidylate synthase cDNA:	
JOURNAL	phylogenetic comparison with human and mouse thymidylate synthases	
MEDLINE	Biochim. Biophys. Acta 1261 (2), 233-242 (1995)	
PUBMED	95226450	
FEATURES	7711067	
source	Location/Qualifiers	
	1..1598	
	/organism="Rattus norvegicus"	
	/strain="Reuber"	
	/db_xref="taxon:10116"	
	/cell_line="H35"	
	/cell_types="hepatoma"	
	1..924	
	/EC number="2.1.1.45"	
	/function="conversion of dUMP to dTMP"	
	/note="dimer; putative"	
	/codon_start=1	
	/product="thymidylate synthase"	
	/protein_id="AAA92340.1"	
	/db_xref="GI:207328"	
	/translation="MLVEGSELQGAQPRTEAPQHGELQVLRQVHIMRCGFKKEDR	
	TGTGILSVFGQARYSLRDEPPLITTKVFNKGLVLELWFKIGNSTAKELSSKGVRI	
	WDANGSRDFLDSGLFSARQGDGPGVYQFMRHFGADYKMDSDYSQGVQDLQKVID	
	TIKTPDPRRIICAMPKDLPLMPCALCOFYVNGSELSCQIYKSGDMGLGVP	
	FNIAGYALLTVMIAHITGLQPGDFVHTLGDHAIYVNIHIEPLKIQLQREPRPFKLRL	
	RKVTIDDFKVEDPQIRGYNPHPTIKEMAV"	
BASE COUNT	438 a 328 c 375 g 457 t	
ORIGIN		
Query Match	45.4%; Score 697.8; DB 10; Length 1598;	
Best Local Similarity	87.7%; Pred. No. 2.4e-134;	
Matches	762; Conservative 0; Mismatches 107; Indels 0; Gaps 0;	
QY	181 CCGCCGACGCGGAGCTGCAGTACCTGGGCGAGATCCACACATCTCGCTCGCGGCTC 240	
Db	58 CCGCAGCATCGAGAACTCCAGTACCTGAGGCAGGTGAGCACATTATGCGCTCGCGCTC 117	
QY	241 AGAAGAGACGACGCGACCGGACCCCTGTCGTATTCGGCATGCGACGCGGCTAC 300	
Db	118 AAGAAGAGACCGCACCGGCACCTGGCACCTGTCGGTGTTCGGCATGCGACGCGTAC 177	
QY	301 AGCTGAGAGATGAATTCCTCTGCTGACACCAACGCTGTTCTGGAAGGCTGTTTG 360	
Db	178 AGCTGAGAGATGAATTCCTCTGCTGACACCAACGAGTATTCGGAAGGCTGTTTG 237	
QY	361 GAGGAGTTGCTGTGTTTATCAAGGATCCACAAATGCTAAAGAGCTGTTCTCCAGGGA 420	
Db	238 GAGGAGTTGTTGTGTTTATTAAGGCTCCCAATGCTTAAGAGTCTCTCCAGGGA 297	
QY	421 GTGAAATCTGGGATGCCAATGGATCCCGAGACTTTTGGACAGCCCTGGGATTCCTACC 480	
Db	298 GTGGAATTTGGATGCCAATGGTCCCGAGACTTTTGGACAGCTTGGATTCCTGCC 357	
QY	481 AGAGAAGAAGGGACTTGGGCCAGTTTATGGCTTCCAGTGGAGGCAATTTTGGGCGAGAA 540	
Db	358 CGACAGAGAGAGACCTGGGCCAGTTTATGGATTCCAGTGGAGACATTTTGGAGCAGAC 417	
QY	541 TACAGAGATAGGAATCAGATATTTCAGACAGGAGTTGACCACTGCAAGAGATGATT 600	
Db	418 TACAAAGATAGGATTCAGATTACTCGGTCAGAGGAGTACAGCAGCTGCACAAAGTATT 477	
QY	601 GACACCATCAAAACCAACCTCGACGACGAGAGATCATGTCGCCCTGGGATCCAGA 660	
Db	478 GACACCATCAAAACCAACCCCGATGACAGAGAGATCATGTCGCTGCGTGGAAACCCAAA 537	
QY	661 GATCTTCTCTGATGCGCTGCTCCATGCGCTCTGCGCTCTGCGCTCTGATGTTGTTGAAC 720	
Db	538 GATCTTCTCTGATGCGCTGCTCTCTGCGCTCTGCGCTCTGCTCTGCTCTGTTGTTGTTG 597	
QY	721 AGTGAGCTGTCTGCGCTGCTGATGCGGAGATCGGGAGACATCGGCTCGGTGCTCTTC 780	
Db	598 GGGGAGCTGTCTGCGCTGCTTACAGCGGTGAGAGATATGGTCTGGGTGTCGCTTC 657	
QY	781 AACATCGCCAGCTGACGCTCTGCTGATGCGGAGATCGGATGCGGCTCGGCTCGGCTCG 840	
Db	658 AACATGCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 717	
QY	841 GGTGACTTATACACTTTGGGAGATGCGACATATTACCTGAATCAGATCGAGCGACTG 900	
Db	718 GGTGATTTTGTCCATCTTTGGGAGATGCGACATTTATCTGAATCATATTGAGCCACTG 777	
QY	901 AAAATTCAGCTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGGATTCCTTCGAAAAGTT 960	
Db	778 AAAATTCAGCTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGGATTCCTTCGAAAAGTT 837	
QY	961 GAGAAATTCAGTACCTTCAAGCTGCAAGACTTTTCAGATTGAAGGCTACAAATCCCATCCA 1020	
Db	838 GAGCAATTCAGTACCTTCAAGCTGCAAGACTTTTCAGATTGAAGGCTACAAATCCCATCCA 897	
QY	1021 ACTATTAAATGGAATGCTGTTTGGG 1049	
Db	898 ACGATTAAATGGAATGCTGTTTGGG 926	
RESULT 15		
MUSPSTA	MUSPSTA	1349 bp mRNA linear ROD 27-APR-1993
LOCUS	Mouse thymidylate synthase pseudogene mRNA, complete cds.	
DEFINITION		
ACCESSION	M30774.1	
VERSION	GI:200561	
KEYWORDS	pseudogene; thymidylate synthase.	
SOURCE	Mouse LUI3-7 and 3T6 cell line DNA.	
ORGANISM	Mus musculus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	1 (bases 1 to 1349)	
AUTHORS	Li,D.W. and Johnson,L.F.	
TITLE	A mouse thymidylate synthase pseudogene derived from an aberrantly processed RNA molecule	
JOURNAL	Gene 82 (2), 363-370 (1989)	
MEDLINE	90060828	
PUBMED	2573561	
FEATURES	Location/Qualifiers	
	1..1349	
	/organism="Mus musculus"	
	/db_xref="taxon:10090"	
	177..1097	
	/note="pseudo-thymidylate synthase"	
	/pseudo	
	/codon_start=1	
BASE COUNT	345 a 320 c 335 g 349 t	
ORIGIN	1 bp upstream of HindIII site.	
Query Match	41.3%; Score 634.8; DB 10; Length 1349;	
Best Local Similarity	84.6%; Pred. No. 2.8e-121;	
Matches	767; Conservative 0; Mismatches 118; Indels 22; Gaps 4;	
QY	181 CCGCCGACGCGGAGCTGAGTACCTGGGCGAGATCCACACATCTCTCGCTCGCGGCTC 240	
Db	225 CCGGGGATGAGAACTCCAGTACCTGAGGAGGTGAGGACATTTTGTCTCGCGCTC 284	
QY	241 AGAAGAGAGACCCGACGCGCACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300	
Db	285 AAGAAGAGAGACCCGACGCGCACCTGGCACCTGTAGGTGTTCGGCATGCGAGCGATCC 344	
QY	301 AGCTGAGAGATGAATTCCTCTGCTGACACCAACGCTGTTCTGGAAGGCTGTTTG 360	
Db	345 AGCTGAGAGATGAATTCCTCTGCTGACACCAACGCTGTTCTGGAAGGCTGTTTG 403	

```
QY 666 TCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 725
Db 592 TCCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 651
QY 726 GCTCTCTGCGAGCTGTACACAGATCGGAGACATGGGCGCTGGTGTGCTTTCAACAT 785
Db 652 ACTGTCTTGGCGCTTACACAGGCTCAGAGATGAGATGAGTGGTCTGGGCGTGGCTTCAACAT 711
QY 786 CGCAGCTACGCCCTCTCTCACATACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 845
Db 712 TGCCAGCTATGCTCTGCTCACTACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 771
QY 846 CTTTATACACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 905
Db 772 TTTTGTCCACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 831
QY 906 TCAGCTTCAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAA 965
Db 832 TCAGCTACAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAC 891
QY 966 AATGTGATCACTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 1025
Db 892 AATGTGATCACTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 951
QY 1026 TAAATGCAATGGCTGTTTA 1046
Db 952 TAAATGCAATGGCTGTTTA 972

RESULT 13
MUSTHSM 961 bp mRNA linear ROD 27-APR-1993
LOCUS Mouse thymidylate synthase mRNA, complete cds.
DEFINITION
ACCESSION M13019
VERSION M13019.1 GI:202029
KEYWORDS thymidylate synthetase.
SOURCE Mouse, cDNA to mRNA, clone pMTS-3.
ORGANISM Mus musculus
REFERENCE
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (Bases 1 to 961)
Perryman,S.M., Rossana,C., Deng,T.L., Vanin,E.F. and Johnson,L.F.
Sequence of a cDNA for mouse thymidylate synthase reveals striking
similarity with the prokaryotic enzyme
Mol. Biol. Evol. 3 (4), 313-321 (1986)
JOURNAL 88174353
MEDLINE 3444407
PUBMED
COMMENT Draft entry and printed copy of the sequence [1] kindly provided by
L.F. Johnson, 29-AUG-1986.
FEATURES
source
1. .961
/organism="Mus musculus"
/db xref="taxon:10090"
<1. .961
/product="this mRNA"
38. .961
/notes="thymidylate synthase (EC 2.1.1.45)"
/codon start=1
/protein id="AAA40439.1"
/db xref="GI:202030"
/translation="MLVYSELSQDAOLSAEPHGELOYLRQVEHLIRGFKEDR
TGUTLSVFGMQARLSLRDEPFLTTKRVFKWGLVLELLNFKGSTNAKELSSKGVRI
WDANGSRDFSLGSADEGLDGLVYGFQWRHFHFGAAYKDMDSYDGGVDDLOKQVID
TITNPDRIITMCANPKOLFMAIPFCHALCQFYVNGELSCQLYORSGDMLGVP
FNIASVALLYMTAHTILOPQDFVHTLGDHLYINHIPLIKIQLQREPRPPFKLIL
RKVETIDDFKVEDFQEGYNPHPTIKMEMAV"
BASE COUNT 253 a 224 c 251 g 233 t
ORIGIN 63 bp upstream of PstI site.
Query Match 46.0%; Score 706.6; DB 10; Length 961;
Best Local Similarity 85.5%; Pred. No. 3.4e-136;
Matches 787; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
```

RESULT 14
RATTHYSYN
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS

RATTHYSYN 1598 bp mRNA linear ROD 30-MAY-2000
Rattus norvegicus thymidylate synthase mRNA, complete cds.
L12138
L12138.1 GI:207327
thymidylate synthase.

1026 TAAATGCAATGGCTGTTTA 1046
940 TAAATGCAATGGCTGTTTA 960

QY 1026 TAAATGCAATGGCTGTTTA 1046
Db 940 TAAATGCAATGGCTGTTTA 960

QY 906 TCAGCTTCAGGAGAACCCAGACCTTTCCCAAGCTCAGGATTTCTCGAAAAGTTGAGAA 965
Db 820 TCAGCTTCAGGAGAACCCAGACCTTTCCCAAGCTCAGGATTTCTCGAAAAGTTGAGAC 879

QY 966 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 1025
Db 880 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 939

QY 846 CTTTATACACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 905
Db 760 TTTTGTCCACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 819

QY 906 TCAGCTTCAGGAGAACCCAGACCTTTCCCAAGCTCAGGATTTCTCGAAAAGTTGAGAA 965
Db 820 TCAGCTTCAGGAGAACCCAGACCTTTCCCAAGCTCAGGATTTCTCGAAAAGTTGAGAC 879

QY 966 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 1025
Db 880 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 939

QY 846 CTTTATACACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 905
Db 760 TTTTGTCCACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 819

QY 906 TCAGCTTCAGGAGAACCCAGACCTTTCCCAAGCTCAGGATTTCTCGAAGAGTTGAGAA 965
Db 820 TCAGCTTCAGGAGAACCCAGACCTTTCCCAAGCTCAGGATTTCTCGAAAAGTTGAGAC 879

QY 966 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 1025
Db 892 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 951

QY 1026 TAAATGCAATGGCTGTTTA 1046
Db 952 TAAATGCAATGGCTGTTTA 972

QY 666 TCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 725
Db 592 TCCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 651

QY 726 GCTCTCTGCGAGCTGTACACAGATCGGAGACATGGGCGCTGGTGTGCTTTCAACAT 785
Db 652 ACTGTCTTGGCGCTTACACAGGCTCAGAGATGAGATGAGTGGTCTGGGCGTGGCTTCAACAT 711

QY 786 CGCAGCTACGCCCTCTCTCACATACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 845
Db 712 TGCCAGCTATGCTCTGCTCACTACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 771

QY 846 CTTTATACACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 905
Db 772 TTTTGTCCACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 831

QY 906 TCAGCTTCAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAA 965
Db 832 TCAGCTACAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAC 891

QY 966 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 1025
Db 892 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 951

QY 1026 TAAATGCAATGGCTGTTTA 1046
Db 952 TAAATGCAATGGCTGTTTA 972

QY 666 TCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 725
Db 592 TCCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 651

QY 726 GCTCTCTGCGAGCTGTACACAGATCGGAGACATGGGCGCTGGTGTGCTTTCAACAT 785
Db 652 ACTGTCTTGGCGCTTACACAGGCTCAGAGATGAGATGAGTGGTCTGGGCGTGGCTTCAACAT 711

QY 786 CGCAGCTACGCCCTCTCTCACATACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 845
Db 712 TGCCAGCTATGCTCTGCTCACTACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 771

QY 846 CTTTATACACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 905
Db 772 TTTTGTCCACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 831

QY 906 TCAGCTTCAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAA 965
Db 832 TCAGCTACAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAC 891

QY 966 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 1025
Db 892 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 951

QY 1026 TAAATGCAATGGCTGTTTA 1046
Db 952 TAAATGCAATGGCTGTTTA 972

QY 666 TCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 725
Db 592 TCCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 651

QY 726 GCTCTCTGCGAGCTGTACACAGATCGGAGACATGGGCGCTGGTGTGCTTTCAACAT 785
Db 652 ACTGTCTTGGCGCTTACACAGGCTCAGAGATGAGATGAGTGGTCTGGGCGTGGCTTCAACAT 711

QY 786 CGCAGCTACGCCCTCTCTCACATACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 845
Db 712 TGCCAGCTATGCTCTGCTCACTACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 771

QY 846 CTTTATACACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 905
Db 772 TTTTGTCCACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 831

QY 906 TCAGCTTCAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAA 965
Db 832 TCAGCTACAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAC 891

QY 966 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 1025
Db 892 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 951

QY 1026 TAAATGCAATGGCTGTTTA 1046
Db 952 TAAATGCAATGGCTGTTTA 972

QY 666 TCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 725
Db 592 TCCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 651

QY 726 GCTCTCTGCGAGCTGTACACAGATCGGAGACATGGGCGCTGGTGTGCTTTCAACAT 785
Db 652 ACTGTCTTGGCGCTTACACAGGCTCAGAGATGAGATGAGTGGTCTGGGCGTGGCTTCAACAT 711

QY 786 CGCAGCTACGCCCTCTCTCACATACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 845
Db 712 TGCCAGCTATGCTCTGCTCACTACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 771

QY 846 CTTTATACACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 905
Db 772 TTTTGTCCACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 831

QY 906 TCAGCTTCAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAA 965
Db 832 TCAGCTACAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAC 891

QY 966 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 1025
Db 892 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 951

QY 1026 TAAATGCAATGGCTGTTTA 1046
Db 952 TAAATGCAATGGCTGTTTA 972

QY 666 TCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 725
Db 592 TCCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 651

QY 726 GCTCTCTGCGAGCTGTACACAGATCGGAGACATGGGCGCTGGTGTGCTTTCAACAT 785
Db 652 ACTGTCTTGGCGCTTACACAGGCTCAGAGATGAGATGAGTGGTCTGGGCGTGGCTTCAACAT 711

QY 786 CGCAGCTACGCCCTCTCTCACATACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 845
Db 712 TGCCAGCTATGCTCTGCTCACTACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 771

QY 846 CTTTATACACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 905
Db 772 TTTTGTCCACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 831

QY 906 TCAGCTTCAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAA 965
Db 832 TCAGCTACAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAC 891

QY 966 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 1025
Db 892 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 951

QY 1026 TAAATGCAATGGCTGTTTA 1046
Db 952 TAAATGCAATGGCTGTTTA 972

QY 666 TCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 725
Db 592 TCCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 651

QY 726 GCTCTCTGCGAGCTGTACACAGATCGGAGACATGGGCGCTGGTGTGCTTTCAACAT 785
Db 652 ACTGTCTTGGCGCTTACACAGGCTCAGAGATGAGATGAGTGGTCTGGGCGTGGCTTCAACAT 711

QY 786 CGCAGCTACGCCCTCTCTCACATACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 845
Db 712 TGCCAGCTATGCTCTGCTCACTACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 771

QY 846 CTTTATACACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 905
Db 772 TTTTGTCCACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 831

QY 906 TCAGCTTCAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAA 965
Db 832 TCAGCTACAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAC 891

QY 966 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 1025
Db 892 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 951

QY 1026 TAAATGCAATGGCTGTTTA 1046
Db 952 TAAATGCAATGGCTGTTTA 972

QY 666 TCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 725
Db 592 TCCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 651

QY 726 GCTCTCTGCGAGCTGTACACAGATCGGAGACATGGGCGCTGGTGTGCTTTCAACAT 785
Db 652 ACTGTCTTGGCGCTTACACAGGCTCAGAGATGAGATGAGTGGTCTGGGCGTGGCTTCAACAT 711

QY 786 CGCAGCTACGCCCTCTCTCACATACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 845
Db 712 TGCCAGCTATGCTCTGCTCACTACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 771

QY 846 CTTTATACACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 905
Db 772 TTTTGTCCACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 831

QY 906 TCAGCTTCAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAA 965
Db 832 TCAGCTACAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAC 891

QY 966 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 1025
Db 892 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 951

QY 1026 TAAATGCAATGGCTGTTTA 1046
Db 952 TAAATGCAATGGCTGTTTA 972

QY 666 TCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 725
Db 592 TCCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 651

QY 726 GCTCTCTGCGAGCTGTACACAGATCGGAGACATGGGCGCTGGTGTGCTTTCAACAT 785
Db 652 ACTGTCTTGGCGCTTACACAGGCTCAGAGATGAGATGAGTGGTCTGGGCGTGGCTTCAACAT 711

QY 786 CGCAGCTACGCCCTCTCTCACATACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 845
Db 712 TGCCAGCTATGCTCTGCTCACTACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 771

QY 846 CTTTATACACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 905
Db 772 TTTTGTCCACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 831

QY 906 TCAGCTTCAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAA 965
Db 832 TCAGCTACAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAC 891

QY 966 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 1025
Db 892 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 951

QY 1026 TAAATGCAATGGCTGTTTA 1046
Db 952 TAAATGCAATGGCTGTTTA 972

QY 666 TCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 725
Db 592 TCCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 651

QY 726 GCTCTCTGCGAGCTGTACACAGATCGGAGACATGGGCGCTGGTGTGCTTTCAACAT 785
Db 652 ACTGTCTTGGCGCTTACACAGGCTCAGAGATGAGATGAGTGGTCTGGGCGTGGCTTCAACAT 711

QY 786 CGCAGCTACGCCCTCTCTCACATACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 845
Db 712 TGCCAGCTATGCTCTGCTCACTACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 771

QY 846 CTTTATACACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 905
Db 772 TTTTGTCCACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 831

QY 906 TCAGCTTCAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAA 965
Db 832 TCAGCTACAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAC 891

QY 966 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 1025
Db 892 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 951

QY 1026 TAAATGCAATGGCTGTTTA 1046
Db 952 TAAATGCAATGGCTGTTTA 972

QY 666 TCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 725
Db 592 TCCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 651

QY 726 GCTCTCTGCGAGCTGTACACAGATCGGAGACATGGGCGCTGGTGTGCTTTCAACAT 785
Db 652 ACTGTCTTGGCGCTTACACAGGCTCAGAGATGAGATGAGTGGTCTGGGCGTGGCTTCAACAT 711

QY 786 CGCAGCTACGCCCTCTCTCACATACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 845
Db 712 TGCCAGCTATGCTCTGCTCACTACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 771

QY 846 CTTTATACACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 905
Db 772 TTTTGTCCACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 831

QY 906 TCAGCTTCAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAA 965
Db 832 TCAGCTACAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAC 891

QY 966 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 1025
Db 892 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 951

QY 1026 TAAATGCAATGGCTGTTTA 1046
Db 952 TAAATGCAATGGCTGTTTA 972

QY 666 TCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 725
Db 592 TCCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 651

QY 726 GCTCTCTGCGAGCTGTACACAGATCGGAGACATGGGCGCTGGTGTGCTTTCAACAT 785
Db 652 ACTGTCTTGGCGCTTACACAGGCTCAGAGATGAGATGAGTGGTCTGGGCGTGGCTTCAACAT 711

QY 786 CGCAGCTACGCCCTCTCTCACATACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 845
Db 712 TGCCAGCTATGCTCTGCTCACTACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 771

QY 846 CTTTATACACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 905
Db 772 TTTTGTCCACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 831

QY 906 TCAGCTTCAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAA 965
Db 832 TCAGCTACAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAC 891

QY 966 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 1025
Db 892 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 951

QY 1026 TAAATGCAATGGCTGTTTA 1046
Db 952 TAAATGCAATGGCTGTTTA 972

QY 666 TCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 725
Db 592 TCCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 651

QY 726 GCTCTCTGCGAGCTGTACACAGATCGGAGACATGGGCGCTGGTGTGCTTTCAACAT 785
Db 652 ACTGTCTTGGCGCTTACACAGGCTCAGAGATGAGATGAGTGGTCTGGGCGTGGCTTCAACAT 711

QY 786 CGCAGCTACGCCCTCTCTCACATACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 845
Db 712 TGCCAGCTATGCTCTGCTCACTACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 771

QY 846 CTTTATACACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 905
Db 772 TTTTGTCCACATCTTGGGAGATGCGATATTTACCTGATTCATCGAGCCACTGAAAT 831

QY 906 TCAGCTTCAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAA 965
Db 832 TCAGCTACAGGAGACCGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAC 891

QY 966 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 1025
Db 892 AATGTGATCACTTCAAGCTCAGAGCTTTTCAGATTGAAGGTTATATCCACATCCACAT 951

QY 1026 TAAATGCAATGGCTGTTTA 1046
Db 952 TAAATGCAATGGCTGTTTA 972

QY 666 TCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 725
Db 592 TCCCTCTGATGGCGCTCCCTCCATGCCATGCCCTCTGCCAGTCTTATGTGTGTAACAGTGA 651

QY 726 GCTCTCTGCGAGCTGTACACAGATCGGAGACATGGGCGCTGGTGTGCTTTCAACAT 785
Db 652 ACTGTCTTGGCGCTTACACAGGCTCAGAGATGAGATGAGTGGTCTGGGCGTGGCTTCAACAT 711

QY 786 CGCAGCTACGCCCTCTCTCACATACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 845
Db 712 TGCCAGCTATGCTCTGCTCACTACATGATTCGCGACATACAGGCGCTGAGACCGAGTGA 771

richards. S., Gibbs, R.A.

Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAP Plate: 36 Row: J Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarly but not
passively to protein.

FEATURES source

```

1. 986
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="CCECH 11"
/clone="MGC:28246 IMAGE:3994204"
/tissue type="Mammary tumor metastasized to lung. Tumor
arose spontaneously from a senescent normal mammary
(cloneal) outgrowth infected with the virus MMTV."
/clone_lib="NCI CGAP_Lu29"
/lab host="DH10B"
/notes="Vector: pCWB-SPORT6"
50. 973
/codon_start=1
/product="Unknown (protein for MGC:28246)"
/protein_id="AAH20139.1"
/db_xref="GI:18044897"
/translacion="MLVYVGLSDAQQLSAPQHGELQYLRVETHLLCGFKKEDR
TGLTSLVGNQARYSALDEPEPLLTPYKVPKWGLELLFWIGSTNAKELSSKGVR1
WDGTSDFLDSIGYSARQEGELGPGYGFQWFGYGAEKMDSDYSGGVQDLQKVID
TTKTNPDRI1ICMANGEDLPLMALPPCHALCFYVYNGEISLCQLYQSGMDGLGVP
FIASYALMTYAHITGLQPGDFVHTLGDAL1YLNIEPLKIQLOQRFPFPFKLLI
RKVETLDDFKVEDFOEGYNVPTIKMEAV"
263. 9
237. 7

```

505

BASE COUNT	270 a	226 c	253 g	237 c
ORIGIN				March 986.

```

Query Match          46.3%; Score 711.4; DB 10;
Best Local Similarity 85.8%; Pred. No. 3.5e-137;
Matches 790; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

```

[illegible]

RESULT 12	ROD 07-AUG-2002
EC020139	linear
	mRNA
	986 bp
	complete cds.

BC020139	986 bp	IMAGE:3994204, mRNA, complete cds.
Mus musculus, clone MGC:28246		
BC020139		
BC020139.1	GI:18044896	
MGC.		
house mouse.		
Mus musculus		
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		
1 (bases 1 to 986)		
Strausberg, R.		
Direct Submission		
Submitted (19-DEC-2001)		National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer		
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
USA		
NIH-MGC Project URL: http://mgc.nci.nih.gov		
Contact: MGC help desk		
Email: cgapbs-remail.nih.gov		
Tissue Procurement: Gilbert Smith, Ph.D.		
cDNA Library Preparation: Life Technologies, Inc.		
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
RNA Sequencing by: Baylor College of Medicine Human Genome		
Sequencing Center		
Center code: BCM-HGSC		
Web site: http://www.hgsc.bcm.tmc.edu/cdna/		
Contact: ang@bcm.tmc.edu		
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,		
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,		

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 840
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18p11.32"
/cell_line="KATO III"
1. 840
/gene="TS"
1. 840
/note="TS"
/note="TS alternative splicing variant The protein encoded by this gene is 34aa shorter than the full-length isoform. This variant uses an alternative splice site resulting in it lacking 102nt in the coding region compared to the full-length isoform reported in Acc# D00596"
/codon_start=1
/evidence=experimental
/product="thymidylate synthase"
/protein_id="BAB83677.1"
/db_xref="GI:18150851"
/translation="MPVAGSELPRRPLPPAAQRDAERPPHGELOYLGQIHLRCLVKQDRGTGCTLSVFGQARYSLRDEPLLTITKEVPKGVLELLMFWFKGSKNAKLSKGVKIDWANGSRDPLDSLGFSREGLDGVYGFQWRHFGAEDYRDMESLDPLMALPFCALCQVQNSQSLQYQKSGDMGLGVFNFIASVALLYMIATHITGLKPGQFIHTLGDHAIYHNLHLEPLKIQLQREPRPPKRLRLRKVEKIDDFKAEQFIEGYNPHPTIRKE
NAV"
BASE COUNT 199 a 224 c 234 g 183 t
ORIGIN
Query Match 47.4%; Score 728; DB 9; Length 840;
Best Local Similarity 89.2%; Pred. No. 1.2e-140;
Matches 840; Conservative 0; Mismatches 0; Indels 102; Gaps 1;
QY 106 ATGCCTGTGCCGGCTCGAGAGTCGCGCGCGCGCCCTTCCCGCCGCGCAGGAGCGG 165
Db 1 ATGCCTGTGCCGGCTCGAGAGTCGCGCGCGCGCCCTTCCCGCCGCGCAGGAGCGG 60
QY 166 GAGCCGAGCGCGCTCCCGCGCAGCGGAGCTGCGAGTACTTGGGGCAGATCCAAACATC 225
Db 61 GAGCGCGAGCGCGCTCCCGCGCAGCGGAGCTGCAGTACCTTGGGGCAGATCCAAACATC 120
QY 226 CTCGCTGCGCGCTCAGAGAGGACGCCAGCGGACCCGCGACCCCTGTCGGTATTCGGC 285
Db 121 CTCGCTGCGCGCTCAGAGAGGACGCCAGCGGACCCGCGACCCCTGTCGGTATTCGGC 180
QY 286 ATGCAGCGCGCTACAGCCTGAGAGATGAATTCCTCTGTGTGCAACCAACGCTGTTC 345
Db 181 ATGCAGCGCGCTACAGCCTGAGAGATGAATTCCTCTGTGTGCAACCAACGCTGTTC 240
QY 346 TGAAGCGGTGTTTTGGAGGAGTTCCTGTGTGTTTATCAAGGGATCCCAAAATGCTAAAGAG 405
Db 241 TGAAGCGGTGTTTTGGAGGAGTTCCTGTGTGTTTATCAAGGGATCCCAAAATGCTAAAGAG 300
QY 406 CTGCTTCCAGGAGTGAAAAATCTGGGATGCCAATGGATCCGAGACTTTTGGACAGC 465
Db 301 CTGCTTCCAGGAGTGAAAAATCTGGGATGCCAATGGATCCGAGACTTTTGGACAGC 360


```
QY 1230 TAATAAAGGCTTTGAGTTAACTCACTGAGGGTATCTGCAATGCTGAGGTTATGAACAA 1289
Db 1206 TAATAAAGGCTTTGAGTTAACTCACTGAGGGTATCTGCAATGCTGAGGTTATGAACAA 1265
QY 1290 AGTGAGGAGAAATGAATGATGCTCTTGAAGAAACATGATGCTCAATCCCA 1349
Db 1266 AGTGAGGAGAAATGAATGATGCTCTTGAAGAAACATGATGCTCAATCCCA 1325
QY 1350 CGTACTTATAAAGAGGTTGGTGAATTTCCCAAGCTATTTTGGAAATATTTTAGAATAT 1409
Db 1326 CGTACTTATAAAGAGGTTGGTGAATTTCCCAAGCTATTTTGGAAATATTTTAGAATAT 1385
QY 1410 TTTAAGAAATTTCCCAAGCTATTTCCCTCAATCTGAGGGAGCTGAGTAACACCATCGATCA 1469
Db 1386 TTTAAGAAATTTCCCAAGCTATTTCCCTCAATCTGAGGGAGCTGAGTAACACCATCGATCA 1445
QY 1470 TGATGTAGAGTGGTGTATGAACCTTTTANAGTGTGTTTATGATGCTGATATATAA 1523
Db 1446 TGATGTAGAGTGGTGTATGAACCTTTTANAGTGTGTTTATGATGCTGATATATAA 1505
QY 1524 AGAAGTGTCTGTC 1536
Db 1506 AGAAGTGTCTGTC 1518

RESULT 9
AX069365
LOCUS AX069365 942 bp DNA linear PAT 25-JAN-2001
DEFINITION Sequence 29 from Patent WO0102600.
ACCESSION AX069365
VERSION AX069365.1 GI:12579227
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE Yuan, C.S.
JOURNAL Detection of analytes using attenuated enzymes
GENERAL PATENT: WO 0102600-A 29 11-JAN-2001;
ATOMIC (US)
FEATURES
LOCATION/Qualifiers
SOURCE 1..942
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Human thymidylate synthase gene: exons 1-8"
BASE COUNT 236 a 247 c 257 g 202 t
ORIGIN

Query Match 61.3%; Score 942; DB 6; Length 942;
Best Local Similarity 100.0%; Pred. No. 5e-185;
Matches 942; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 ATGCTGTGCGGCTCGAGGTGCGCGCGCGCTTGGCCCGCCCGCCACAGAGCGG 165
Db 1 ATGCTGTGCGGCTCGAGGTGCGCGCGCGCTTGGCCCGCCCGCCCGCCACAGAGCGG 60
QY 166 GACCGCGAGCGGCTCGCGCGCGAGTGCAGTACCTGGGGGAGATCCACATC 225
Db 61 GACCGCGAGCGGCTCGCGCGCGAGTGCAGTACCTGGGGGAGATCCACATC 120
QY 226 CTCGCTGCGGCTCAGGAGGAGGACCGGACCGGACCGGACCGCTGCGGTATGCGG 285
Db 121 CTCGCTGCGGCTCAGGAGGAGGACCGGACCGGACCGGACCGCTGCGGTATGCGG 180
QY 286 ATGAGGCGGCTACAGCTGAGAGATGAATTCCTCTGCTGACACCAACAGTGTTC 345
Db 181 ATGAGGCGGCTACAGCTGAGAGATGAATTCCTCTGCTGACACCAACAGTGTTC 240
QY 346 TGGAGGGTGTGTTGAGGAGTGTGTTGTTTATCAAGGATCCCAATGCTTAAGAG 405
Db 241 TGGAGGGTGTGTTGAGGAGTGTGTTGTTTATCAAGGATCCCAATGCTTAAGAG 300
```

```
QY 406 CTCTCTTCAAGGGAGTGAATCTGGGATGCAATGATGATCCCGAGACTTTTGGACAGC 465
Db 301 CTCTCTTCAAGGGAGTGAATCTGGGATGCAATGATGATCCCGAGACTTTTGGACAGC 360
QY 466 CTGGGATTTCCACACAGAGAGAGGGGACTTTGGGCCCGAGTTTATGCTTCCAGTGGAGG 525
Db 361 CTGGGATTTCCACACAGAGAGAGGGGACTTTGGGCCCGAGTTTATGCTTCCAGTGGAGG 420
QY 526 CATTTTGGGGGAGAGATACAGAGATATGGAATCAGATTTATCAGGACAGGAGTTGACCAA 585
Db 421 CATTTTGGGGGAGAGATACAGAGATATGGAATCAGATTTATCAGGACAGGAGTTGACCAA 480
QY 586 CTCGAAAGAGTATTTGACACCATCAAAACCCCTTGACGACAGAGATCATCATGTGC 645
Db 481 CTCGAAAGAGTATTTGACACCATCAAAACCCCTTGACGACAGAGATCATCATGTGC 540
QY 646 GCTTGAATCCAAAGAGATTTCTCTGATGGCGCTGCTCCATGCGCTCTCTGCCAG 705
Db 541 GCTTGAATCCAAAGAGATTTCTCTGATGGCGCTGCTCCATGCGCTCTCTGCCAG 600
QY 706 TTCTATGTGTGAACAGTGAAGTCTCTCTGCTGCTGACGATGACGAGATCGGAGATGGGC 765
Db 601 TTCTATGTGTGAACAGTGAAGTCTCTCTGCTGCTGACGATGACGAGATCGGAGATGGGC 660
QY 766 CTGGTGTGCTTTCAACATCGCAGCTAGCCCTGCTCAGTACATGATGCGCACATC 825
Db 661 CTGGTGTGCTTTCAACATCGCAGCTAGCCCTGCTCAGTACATGATGCGCACATC 720
QY 826 ACGGGCTGAGCCAGGTGATCTTTATACACTTTGGGAGATGCACATTTACTGAAT 885
Db 721 ACGGGCTGAGCCAGGTGATCTTTATACACTTTGGGAGATGCACATTTACTGAAT 780
QY 886 CACATCAGCCACTGGAATTCAGCTTCAGCGAAGACCCAGACCTTTCCCAAGCTCAGG 945
Db 781 CACATCAGCCACTGGAATTCAGCTTCAGCGAAGACCCAGACCTTTCCCAAGCTCAGG 840
QY 946 ATTCTTCAAAAGTTCAGAAAATTCAGTACTTCAAGAGCTTCAAGACTTTCAGATTGAAGG 1005
Db 841 ATTCTTCAAAAGTTCAGAAAATTCAGTACTTCAAGAGCTTCAAGACTTTCAGATTGAAGG 900
QY 1006 TACAATCCGATCCAACTATTAAAATGGAATGCTGTTAG 1047
Db 901 TACAATCCGATCCAACTATTAAAATGGAATGCTGTTAG 942

RESULT 10
AX397717
LOCUS AX397717 942 bp DNA linear PAT 18-MAY-2002
DEFINITION Sequence 143 from Patent WO0210366.
ACCESSION AX397717
VERSION AX397717.1 GI:21068438
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE Chen, H., Preimer, N.B. and Novak, T.
JOURNAL Methods and compositions for diagnosing and treating chromosome-18p related disorders
Patent: WO 0210366-A 143 07-FEB-2002;
Millennium Pharmaceuticals, Inc. (US) ; The Regents of The University of California (US)
FEATURES
LOCATION/Qualifiers
SOURCE 1..942
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 236 a 247 c 257 g 202 t
ORIGIN

Query Match 61.3%; Score 942; DB 6; Length 942;
Best Local Similarity 100.0%; Pred. No. 5e-185;
Matches 942; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Db 1340 CGTACTTATAAGAGGTTGGTGAATTTACAGACTATTTTGGATAATTTTGAATAT 1399
QY 1410 TTTAAGAAATTCACAGCTATTTCCCTCAAACTCTGAGGAGCTAGTAACACATCGATCA 1469
Db 1400 TTTAAGAAATTCACAGCTATTTCCCTCAAACTCTGAGGAGCTAGTAACACATCGATCA 1459
QY 1470 TGATGTAGAGTGTGGTATGAACCT-----TTANAGTTGTTTATATATGTGCTATAATAA 1523
Db 1460 TGATGTAGAGTGTGGTATGAACCTTTAAAGTTATAGTTGTTTATATATGTGCTATAATAA 1519
QY 1524 AGAAGTGTCTGC 1536
Db 1520 AGAAGTGTCTGC 1532

RESULT 8
AB062290 1567 bp mRNA linear PRI 23-MAY-2002
LOCUS Homo sapiens OK/SW-cl.29 mRNA for thymidylate synthetase, complete
DEFINITION
ACCESSION AB062290
VERSION AB062290.1 GI:21104405
KEYWORDS
SOURCE Homo sapiens colon adenocarcinoma cell_line:SW620 cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Shichiyo,S. and Itoh,K.
TITLE Identification of immuno-peptidmics that recognized by
tumor-reactive CTL generated from TIL of colon cancer patients
JOURNAL Unpublished
2 (bases 1 to 1567)
AUTHORS Itoh,K. and Shichiyo,S.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-2001) Shigeaki Shichiyo, Kurume University School
of Medicine, Department of Immunology; 67 Asahi-machi, Kurume,
Fukuoka 830-0011, Japan (E-mail:shichiyo@med.kurume-u.ac.jp,
Tel:81-942-31-7551, Fax:81-942-31-7699)
FEATURES
Location/Qualifiers
1..1567
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="SW620"
/cell_type="colon adenocarcinoma"
1..1567
/genes="OK/SW-cl.29"
82..1023
/genes="OK/SW-cl.29"
/codon_start=1
/product="thymidylate synthetase"
/protein_id="BAB93473.1"
/db_xref="GI:21104406"
/translation="MPVAGSELPRPLPFAAQRDAEPHPGELQYLQGTQHLRGG
VRKDRGTGSLVFGMOARYSLRDEPLLTTRVFWKGLVELELWFIKGTSTNAKELS
SKGVKWDANGSDLEFDSLGSSTREBGLGVFGWRFHGAERDMESDYSOGVDQ
LQRVIDTKINPDRLINMCANRPDLPLMALPPCHALCFYVNSELSCOLYORSQ
MGLGVPTIASYALTYTAHTITGLKPGDFIRLGDHLYLHNLHIEPLKIQLOREPRFF
PKLRLKFKVKIDDFKAEQDFQEGYNPHPTIKMEAV"
BASE COUNT 431 a 366 c 389 g 381 t
ORIGIN
Query Match 95.9%; Score 1473.6; DB 9; Length 1567;
Best Local Similarity 98.8%; Pred. No. 2.7e-295;
Matches 1495; Conservative 0; Mismatches 12; Indels 6; Gaps 1;
QY 30 CGTCCGCGCGCCACATGGCGCTGCCTCCGTCGCGCGGCACATTCGCCTGCCTCGCTC 89
Db 6 CCTGCCACCGCGCCACATGGCGCTGCCTCCGTCGCGCGGCACATTCGCCTGCCTCGCTC 65
QY 90 CCGCGCGCGCGCGCCACATGGCGCTGCCTCCGCGCGGCACATTCGCCTGCCTCGCTC 149
Db 66 CCGCGCGCGCGCGCCACATGGCGCTGCCTCCGCGCGGCACATTCGCCTGCCTCGCTC 125
```

SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1569)
 TITLE Direct Submission
 AUTHORS Strausberg, R.
 JOURNAL Submitted (07-SEP-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amandansystemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
 Greene, Mark Kettman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 31 Row: 0 Column: 4
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Similarity but not
 identity to protein

FEATURES
 Location/Qualifiers
 1..1569
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:22884 IMAGE:4048625"
 /tissue_type="Bone marrow, chronic myelogenous leukemia"
 /clone_lib="NIH MGC 54"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"
 96..1037
 /codon_start=1
 /product="Similar to thymidylate synthetase"
 /protein_id="AAH1391.1"
 /db_xref="GI:15530267"
 /translation="MPVAGSELPRRPLPPAAQERDPRPHGELQYLOIHLIRCG
 VRKDRGTGTLVFGMQARYSLRDEFPLLTKVFWKVLLELLWFKIGSTNAKELS
 SKGVKWDANGSRDLDSFGSTREBGLDPVYQWRHFGARYRMESDVGQGVQD
 LQRVIDTKNPDRIINCAWNPDLPLMPCFALCOFYVNSLESCQYQKSGD
 MGLGVFNFIASVALLVYIAHITGLKPGDFHTLGLDAHIYLNHIEPLKIQLOREPRP
 PKLRLKVEKIDDFKAEQDFQIEGVNPHPTIKMEMAV"
 BASE COUNT 425 a 368 c 398 g 378 t
 ORIGIN

Query Match 96.6%; Score 1483.2; DB 9; Length 1569;
 Best Local Similarity 99.2%; Pred. No. 2.8e-297;
 Matches 1501; Conservative 0; Mismatches 6; Indels 6; Gaps 1;

QY 30 CGTCCGCGCGCCACTTGGCTGCTCGCTCCGCGCGCCACTTGGCTCCGCTCCGTC 89
 Db 20 CTTGCCACCGCGCCACTTGGCTGCTCGCTCCGCGCGCCACTTGGCTCCGCTCCGTC 79

QY 90 CCGCGCGCGCGCGCCACTTGGCTGCTCGCTCCGCGCGCCACTTGGCTCCGCTCCGTC 149
 Db 80 CCGCGCGCGCGCGCCACTTGGCTGCTCGCTCCGCGCGCCACTTGGCTCCGCTCCGTC 139

QY 150 CGCGGACAGGAGCGGAGCGCGCGCGCTGCGAGCTGCGCGCGCGCGCGCTGCGAGTCTGGG 209
 Db 140 CGCGGACAGGAGCGGAGCGCGCGCGCTGCGAGCTGCGCGCGCGCGCGCTGCGAGTCTGGG 199

QY 210 GCAGATCCACACATCTCCGCTCGCGCGCTGCGAGCTGCGCGCGCGCGCGCTGCGAGTCTGGG 269
 Db 200 GCAGATCCACACATCTCCGCTCGCGCGCTGCGAGCTGCGCGCGCGCGCGCTGCGAGTCTGGG 259

QY 270 CCTGTGGTATTCGGCATCGAGCGCGCTACAGCTTCAGAGATGATTCCTCTGCTGAC 329

Db 260 CCTGTGGTATTCGGCATCGAGCGCGCTACAGCTTCAGAGATGATTCCTCTGCTGAC 319
 QY 330 AACCAACCGTGTGTTCTGGAAGGGTGTGTTGGAGAGTGTGCTGTTTATCAAGGGATC 389
 Db 320 AACCAACCGTGTGTTCTGGAAGGGTGTGTTGGAGAGTGTGCTGTTTATCAAGGGATC 379
 QY 390 CACAAATGCTAAAGAGCTGTCTTCCAAAGGAGTGAATCTGGGATGCGCAATGATCCCG 449
 Db 380 CACAAATGCTAAAGAGCTGTCTTCCAAAGGAGTGAATCTGGGATGCGCAATGATCCCG 439
 QY 450 AGACTTTTGGACAGCTGGGATTCCTCCACAGAGAGAGGGGACTTTGGGCCAGTTTA 509
 Db 440 AGACTTTTGGACAGCTGGGATTCCTCCACAGAGAGAGGGGACTTTGGGCCAGTTTA 499
 QY 510 TGGCTTCCAGTGGAGGCAATTTGGGGCAGAAATACAGAGATATGGAATCAGATTATTTCAGG 569
 Db 500 TGGCTTCCAGTGGAGGCAATTTGGGGCAGAAATACAGAGATATGGAATCAGATTATTTCAGG 559
 QY 570 ACAGGAGTTGACCAACTGCAAGAGTGAATGACCAATCAAAACCAACCTGACGACAG 629
 Db 560 ACAGGAGTTGACCAACTGCAAGAGTGAATGACCAATCAAAACCAACCTGACGACAG 619
 QY 630 AAGATCATCATGTGGCTTGGATGCCAAGAGATCTTCTGTGATGGCGCTGCTCCATG 589
 Db 620 AAGATCATCATGTGGCTTGGATGCCAAGAGATCTTCTGTGATGGCGCTGCTCCATG 679
 QY 690 CCATGCCCTCTGCCAGTTCATGTGTGAGTGAACAGTGAAGCTGCTCCAGAGTGTACAGAG 749
 Db 680 CCATGCCCTCTGCCAGTTCATGTGTGAGTGAACAGTGAAGCTGCTCCAGAGTGTACAGAG 739
 QY 750 ATCGGAGAGATGGCGCTGGCTGTCCTTTCAACATCGCAGCTACGCCCTGCTCAGTA 809
 Db 740 ATCGGAGAGATGGCGCTGGCTGTCCTTTCAACATCGCAGCTACGCCCTGCTCAGTA 799
 QY 810 CATGATTGGCCACATCAACGGGCTGAAGCCAGGTGACTTTATACACACTTTGGGAGATGC 869
 Db 800 CATGATTGGCCACATCAACGGGCTGAAGCCAGGTGACTTTATACACACTTTGGGAGATGC 859
 QY 870 ACATATTACTGATCATCATGAGCCACTGAAATTCAGCTTCAGCGAGAACCCAGACC 929
 Db 860 ACATATTACTGATCATCATGAGCCACTGAAATTCAGCTTCAGCGAGAACCCAGACC 919
 QY 930 TTTCCCAAGCTCAGGATTCCTGAAAGTTGAGAAATATGATGATCTCAAGCTGAAGA 989
 Db 920 TTTCCCAAGCTCAGGATTCCTGAAAGTTGAGAAATATGATGATCTCAAGCTGAAGA 979
 QY 990 CTTTCAGATTGAAGGTACATCCCGATCCCACTTATAAATGGAATGGCTGTTTAGGG 1049
 Db 980 CTTTCAGATTGAAGGTACATCCCGATCCCACTTATAAATGGAATGGCTGTTTAGGG 1039
 QY 1050 TGCCTTCAAGAGGCTGGAAGGATATGTCAGTCTTTAGGGTTGGGCTGGATGCCGAGG 1109
 Db 1040 TGCCTTCAAGAGGCTGGAAGGATATGTCAGTCTTTAGGGTTGGGCTGGATGCCGAGG 1099
 QY 1110 TAAAGTTCTTTTGTCTTAAAGAAAGAACTAGGTCAAAATCTGTCGCTGACCTA 1169
 Db 1100 TAAAGTTCTTTTGTCTTAAAGAAAGAAAGAACTAGGTCAAAATCTGTCGCTGACCTA 1159
 QY 1170 TCAGTTATTAATTTTAAAGATGTTGCCATCGGCAAAATGTAATGTCGCGAGTCTTTCCA 1229
 Db 1160 TCAGTTATTAATTTTAAAGATGTTGCCATCGGCAAAATGTAATGTCGCGAGTCTTTCCA 1219
 QY 1230 TAAATAAGGCTTTGAGTTAACTCACTGAGGTTATCGAATGCTGAGGTTATGAACA 1289
 Db 1220 TAAATAAGGCTTTGAGTTAACTCACTGAGGTTATCGAATGCTGAGGTTATGAACA 1279
 QY 1290 AGTGAGGAGAAATGAAATGATGCTCTTACCAAAAACATGATGTCATTTCAATCCCA 1349
 Db 1280 AGTGAGGAGAAATGAAATGATGCTCTTACCAAAAACATGATGTCATTTCAATCCCA 1339
 QY 1350 CCGTATTAAGAAGGTTGGTGAATTTCAACAGTATTTTTCGAAATATTTTAGATAT 1409

```

Db 1508 C 1508

RESULT 6
AX281791
LOCUS
DEFINITION Sequence 200 from Patent WO0177389.
ACCESSION AX281791
VERSION AX281791.1 GI:16609042
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
REFERENCE
AUTHORS Shifman,D., Somogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
Mikita,T. and Tai,J.
TITLE Genes expressed in foam cell differentiation
JOURNAL Patent: WO 0177389-A 200 18-OCT-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
Location/Qualifiers
1..3298
/organism="Homo sapiens"
/db_xref="taxon:9606"
/notes="Incyte ID No: 247220.15"
BASE COUNT 898 a 709 c 744 g 919 t 28 others
ORIGIN
Query Match 97.2%; Score 1493.4; DB 6; Length 3298;
Best Local Similarity 99.3%; Pred. No. 2.3e-299;
Matches 1510; Conservative 0; Mismatches 5; Indels 6; Gaps 1;

QY 22 CTGCTCCGTCGCGCGCGCATTGGCTGCTCCGTCGCTCCGCGCGCGCATTGGCTG 81
Db 1 CTTGCTCCGTCGCGCGCGCATTGGCTGCTCCGTCGCTCCGCGCGCGCATTGGCTG 60
QY 82 CTTGCTCCGTCGCGCGCGCATTGGCTGCTCCGTCGCTCCGCGCGCGCATTGGCTG 141
Db 61 CTTGCTCCGTCGCGCGCGCATTGGCTGCTCCGTCGCTCCGCGCGCGCATTGGCTG 120
QY 142 TTGCCCCCGCGCGCAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 201
Db 121 TTGCCCCCGCGCGCAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 202 TACCTGGGCGAGATCAACATCTCTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCG 261
Db 181 TACCTGGGCGAGATCAACATCTCTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCG 240
QY 262 ACCGGCACCTGTGCGTATTGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 321
Db 241 ACCGGCACCTGTGCGTATTGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 322 CTGCTGACACCAACCGTGTCTGGAAGGGTGTGGAAGGGTGTGGAAGGGTGTGGA 381
Db 301 CTGCTGACACCAACCGTGTCTGGAAGGGTGTGGAAGGGTGTGGAAGGGTGTGGA 360
QY 382 AGGGATCCAAATGCTAAAGAGCTGTCTTCAAGGAGTCAAAATCTGGGATGCCAAT 441
Db 361 AGGGATCCAAATGCTAAAGAGCTGTCTTCAAGGAGTCAAAATCTGGGATGCCAAT 420
QY 442 GGATCCGAGATTTTGGACGCGCTGGGATCTCCACAGAGAGAGAGGGGACTTTGGGC 501
Db 421 GGATCCGAGATTTTGGACGCGCTGGGATCTCCACAGAGAGAGAGGGGACTTTGGGC 480
QY 502 CAGTTTATGCGTCCAGTGAGGCAATTTGGGGCAGAAATACAGAGATATGGAATCAGAT 561
Db 481 CAGTTTATGCGTCCAGTGAGGCAATTTGGGGCAGAAATACAGAGATATGGAATCAGAT 540
QY 562 TATTACGACAGGAGTGTGCAACTGCAAGAGTGTGCAACCATCAAAACCAACCTT 621
Db 541 TATTACGACAGGAGTGTGCAACTGCAAGAGTGTGCAACCATCAAAACCAACCTT 600

```

```

QY 622 GACGACAGAGATCATCATGTGGCTTGGAAATCCAAGAGATCTTCTCTGATGGCGCTG 681
Db 601 GACGACAGAGATCATCATGTGGCTTGGAAATCCAAGAGATCTTCTCTGATGGCGCTG 660
QY 682 CTCCATGCGATGCGCTCTGCGAGTCTATATGTTGTTGAGTGTGAGTGTGAGTGTG 741
Db 661 CTCCATGCGATGCGCTCTGCGAGTCTATATGTTGTTGAGTGTGAGTGTGAGTGTG 720
QY 742 TACGAGAGATCGGAGACATGCGCTCGGTGTCCTTCAACATCGCAGCTACGCGCTG 801
Db 721 TACGAGAGATCGGAGACATGCGCTCGGTGTCCTTCAACATCGCAGCTACGCGCTG 780
QY 802 CTCACGTACATGATTGGCGCATCATCGGCGCTCAAGCCAGCTGACTTTATACACATTG 861
Db 781 CTCACGTACATGATTGGCGCATCATCGGCGCTCAAGCCAGCTGACTTTATACACATTG 840
QY 862 GGAGATGACATATTTACCTGAATCATCATCGAGCCACTGAAATTCAGCTTCAGCGAGAA 921
Db 841 GGAGATGACATATTTACCTGAATCATCATCGAGCCACTGAAATTCAGCTTCAGCGAGAA 900
QY 922 CCAGACCTTTCCCAAGCTCAGGATCTTCGAAAAGTTGAGAAAATTTGATGACTTCAA 981
Db 901 CCAGACCTTTCCCAAGCTCAGGATCTTCGAAAAGTTGAGAAAATTTGATGACTTCAA 960
QY 982 GCTGAGACTTTTCAGATTGAAGGTGACATCCGATCCAACTATTTAAATGGAATGGCT 1041
Db 961 GCTGAGACTTTTCAGATTGAAGGTGACATCCGATCCAACTATTTAAATGGAATGGCT 1020
QY 1042 GTTTAGGTTGCTTTCAAGGAGCTGGAAGATATTGTCACTCTTTAGGGTTGGGCTGGA 1101
Db 1021 GTTTAGGTTGCTTTCAAGGAGCTGGAAGATATTGTCACTCTTTAGGGTTGGGCTGGA 1080
QY 1102 TGCAGAGTAAAGTTCTTTTGTCTTAAAGAAAGAACTAGGTCAAAAATCTGTCC 1161
Db 1081 TGCAGAGTAAAGTTCTTTTGTCTTAAAGAAAGAACTAGGTCAAAAATCTGTCC 1140
QY 1162 GTGACCTATCAGTTATTAAATTTTAAAGATGTTGCCACTGGCAAATGTAATGTGCGAGT 1221
Db 1141 GTGACCTATCAGTTATTAAATTTTAAAGATGTTGCCACTGGCAAATGTAATGTGCGAGT 1200
QY 1222 TCTTTCATATAAAGCTTTGAGTTAACTCACTGAGGGTATCTGACAAATGCTGAGTT 1281
Db 1201 TCTTTCATATAAAGCTTTGAGTTAACTCACTGAGGGTATCTGACAAATGCTGAGTT 1260
QY 1282 ATGAACAAAGTGAGAGAAATGAATGTATCTCTTGAAGCAAAACATGTATGTCATTT 1341
Db 1261 ATGAACAAAGTGAGAGAAATGAATGTATCTCTTGAAGCAAAACATGTATGTCATTT 1320
QY 1342 CAATCCCGTACTTATAAAGAGTTGGTGAATTTCAAGCTATTTTGGAAATATTTT 1401
Db 1321 CAATCCCGTACTTATAAAGAGTTGGTGAATTTCAAGCTATTTTGGAAATATTTT 1380
QY 1402 TAGAATATTTTAAAGATTTCAAGCTATTTCCCTCAAACTCTGAGGGAGCTGAGTAACAC 1461
Db 1381 TAGAATATTTTAAAGATTTCAAGCTATTTCCCTCAAACTCTGAGGGAGCTGAGTAACAC 1440
QY 1462 ATCGATCATGATGTAGAGTGTGTTATGAACCT-----TTANAGTCTTTTATGTTTC 1515
Db 1441 ATCGATCATGATGTAGAGTGTGTTATGAACCTTTAAAGTATAGTCTTTTATGTTTC 1500
QY 1516 TATAATAAAGAGTGTCTGC 1536
Db 1501 TATAATAAAGAGTGTCTGC 1521

```

```

RESULT 7
BC013919
LOCUS
DEFINITION Homo sapiens. Similar to thymidylate synthetase, clone MGC:22884
IMAGE:4048625, mRNA, complete cds.
ACCESSION BC013919
VERSION BC013919.1 GI:15530266
KEYWORDS MGC.

```

Sequencing Center (MISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: misc_mcg@nigri.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, O.L., Masiello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
Tiongson, E.B., Touchman, J.W., Tsurgueon, C., Vogt, J.L., Walker, N.A.,
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILMI at: <http://image.llnl.gov>
Series: IRAL, Place: 4 Row: P Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4507750.

FEATURES

Location/Qualifiers
1..1533
/organism="Homo sapiens"
/db_xref="LocusID:7298"
/db_xref="taxon:9606"
/clone="MGC:1590 IMAGE:3138877"
/tissue type="placenta, choriocarcinoma"
/clone_lib="NIH MGC 21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
78..1019
/codon start=1
/product="thymidylate synthetase"
/protein_id="AAH02567.1"
/db_xref="GI:12803483"
/translation="MPVAGSELPRPLPPAAQEDAEPRPHGLQYLQIHLRCG
VRKDDRTGTLGVFMQARYSLRDEFFLLATKEVFWKGLVLELLFWKINGSTNAKELS
SKGVKIDANGSRDFDLSLGFSTREBGLPVGFWHPHGAERYDMSDYSGGVQD
LQVLDITKINPDRLILMACNPRDLPLMLPCHALCOFVYVNSELSCLYORSQD
MGLVFNFIASVALLFMIAHITGLKPGDFHTLGDHIVLNHIEPLKIQLOREPRP
PKRLRLRKVKIDDFKAEQFQLEGYNPHPTIRMEMAV"

CDS

BASE COUNT 409 a 359 c 389 g 376 t
ORIGIN

Query Match 97.4%; Score 1496.4; DB 9; Length 1533;
Best Local Similarity 99.7%; Pred. No. 5.1e-300;
Matches 1497; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 36 GCGCGCCATTGGCTGCTCGCTCCGCGCGCCACTTTCGCTGCTCGCTCCGCGCGC 95
DB 8 GCGCGCCATTGGCTGCTCGCTCCGCGCGCCACTTTCGCTGCTCGCTCCGCGCGC 67

QY 96 CCGCGCGCCATTGGCTGCTCGCTCCGCGCGCCACTTTCGCTGCTCGCTCCGCGCGC 155
DB 68 CCGCGCGCCATTGGCTGCTCGCTCCGCGCGCCACTTTCGCTGCTCGCTCCGCGCGC 127

QY 156 ACAGAGCGGAGCGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 215
DB 128 ACAGAGCGGAGCGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 187

QY 216 CCACACATCTCCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 275
DB 188 CCACACATCTCCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 247

QY 276 GGTATTTCGGCATGCGCGCGCGCTACGCTGAGAGATGAATTCCTCTGCTGACACCAA 335
DB 248 GGTATTTCGGCATGCGCGCGCGCTACGCTGAGAGATGAATTCCTCTGCTGACACCAA 307

QY 336 ACCTGTGTCTGGAAGGGTGTGTTGGAGAGTGTGTTGGTGTATATCAAGGATCCACAA 395
DB 308 ACCTGTGTCTGGAAGGGTGTGTTGGAGAGTGTGTTGGTGTATATCAAGGATCCACAA 367

QY 396 TGCTAAAGAGCTGTCTTCCAGGAGGTGAATCTGGGATGCCAATGGATCCCGAGCTT 455
DB 368 TGCTAAAGAGCTGTCTTCCAGGAGGTGAATCTGGGATGCCAATGGATCCCGAGCTT 427

QY 456 TTTGGACAGCTTGGGATTCTCCACAGAGAAGAGGAGCTTGGGCCAGTTTATGGCTT 515
DB 428 TTTGGACAGCTTGGGATTCTCCACAGAGAAGAGGAGCTTGGGCCAGTTTATGGCTT 487

QY 516 CAGTGGAGCATTTTGGGGCAGATAACAGAGATATGGAATCAGATTATTCAGGACAGG 575
DB 488 CAGTGGAGCATTTTGGGGCAGATAACAGAGATATGGAATCAGATTATTCAGGACAGG 547

QY 576 AGTTGACCAACTGCAAGAGTGTGACACCATCAACCAACCTCGACGACAGAGAT 635
DB 548 AGTTGACCAACTGCAAGAGTGTGACACCATCAACCAACCTCGACGACAGAGAT 607

QY 636 CATCATGTGCGCTTGGATCCAGAGATCTTCTCTGATGGCGCTCTCATGCCATGC 695
DB 608 CATCATGTGCGCTTGGATCCAGAGATCTTCTCTGATGGCGCTCTCATGCCATGC 667

QY 696 CTTCTCCAGTTTCTATGTGTGGAACAGTGTGCTTCCAGCTGTACAGAGATCGGG 755
DB 668 CTTCTCCAGTTTCTATGTGTGGAACAGTGTGCTTCCAGCTGTACAGAGATCGGG 727

QY 756 AGACATGGCGCTCGGTGTGCTTCAACATCCGAGCTACCGCTGCTGAGTACATGAT 815
DB 728 AGACATGGCGCTCGGTGTGCTTCAACATCCGAGCTACCGCTGCTGAGTACATGAT 787

QY 816 TGGGCACATCAGCGGCTTGAAGCCAGTGTGCTTATACACACTTTGGGAGATGCATAT 875
DB 788 TGGGCACATCAGCGGCTTGAAGCCAGTGTGCTTATACACACTTTGGGAGATGCATAT 847

QY 876 TTACTCTGAATCAATCGAGCCACTGAAATTCAGCTTCAGGAGAACCCAGACCTTCC 935
DB 848 TTACTCTGAATCAATCGAGCCACTGAAATTCAGCTTCAGGAGAACCCAGACCTTCC 907

QY 936 AAAGCTCAGATTTCTGAAAGTGTGAGAAATTCATGATCTTAAAGTGAAGCTTTCA 995
DB 908 AAAGCTCAGATTTCTGAAAGTGTGAGAAATTCATGATCTTAAAGTGAAGCTTTCA 967

QY 996 GATTGAGGGTACATCCGATCCCACTTAAATGAAATGCTGTTTGGGTCCTTT 1055
DB 968 GATTGAGGGTACATCCGATCCCACTTAAATGAAATGCTGTTTGGGTCCTTT 1027

QY 1056 CAAAGAGCTTGAAGGATATTGTGCTTTAGGGGTTGGCTGATGCCAGGTAAAG 1115
DB 1028 CAAAGAGCTTGAAGGATATTGTGCTTTAGGGGTTGGCTGATGCCAGGTAAAG 1087

QY 1116 TTCTTTTTCCTTAAAGAAAGAAAGGAACTAGGTCAAAAATCTGTCGCTGACCTTACGTT 1175
DB 1088 TTCTTTTTCCTTAAAGAAAGAAAGGAACTAGGTCAAAAATCTGTCGCTGACCTTACGTT 1147

QY 1176 ATTAATTTTAAAGATTTGCCACTGGCAAAATGTAACTGTGCCAGTTCTTTCCATAATA 1235
DB 1148 ATTAATTTTAAAGATTTGCCACTGGCAAAATGTAACTGTGCCAGTTCTTTCCATAATA 1207

QY 1236 AAGCTTTGAGTTTAACTCACTGAGGATCTGACAACTGCTGAGGTATGACAAAGTGAG 1295
DB 1208 AAGCTTTGAGTTTAACTCACTGAGGATCTGACAACTGCTGAGGTATGACAAAGTGAG 1267

QY 1296 GAGAATGAAATGTATGCTCTTAGCAAAAAGCACTGATGTGCTTCAATCCCACTACT 1355
DB 1268 GAGAATGAAATGTATGCTCTTAGCAAAAAGCACTGATGTGCTTCAATCCCACTACT 1327

QY 1356 TATAAAGAGGTTTGTGAAATTTCAAGCTATTTTGGAAATTTTGGAAATTTTAAAG 1415
DB 1328 TATAAAGAGGTTTGTGAAATTTCAAGCTATTTTGGAAATTTTGGAAATTTTAAAG 1387

QY 1416 AATTTCAGAGCTATTTCCTCAATCTGAGGAGCTGAGTAACCACTGCATCATGATGT 1475
DB 1388 AATTTCAGAGCTATTTCCTCAATCTGAGGAGCTGAGTAACCACTGCATCATGATGT 1447

QY 1476 AGAGTGTGTTATGAACCTTTTANAGTTGTTTATATGTTGCTTATATAAAGAGTGTCTG 1535
DB 1448 AGAGTGTGTTATGAACCTTTTANAGTTGTTTATATGTTGCTTATATAAAGAGTGTCTG 1507

QY 1536 C 1536

MLGVFPNLIASVALLTYMIAHTIGLKEGDFIHTLGDHAIYILNHIPLKIQLOREPRFF
PKLRIIRKVEKIDDKAEDFOIEGYNPHPTIKMEAV"
1231..1236
/note="pot. polyadenylation signal"
1519..1524
/note="pot. polyadenylation signal"
polya_site
1536
/note="polyadenylation site"
BASE COUNT 390 a 369 c 399 g 378 t
ORIGIN

Query Match 99.8%; Score 1533; DB 9; Length 1536;
Best Local Similarity 99.8%; Pred. No. 1.3e-307;
Matches 1533; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGACACATTGGCTGCTCGCTCCGCGCGCGCCACTTGGCGCTCCCTCCGT 60
Db 1 GGGGGGGGGGACACATTGGCTGCTCGCTCCGCGCGCGCCACTTGGCGCTCCCTCCGT 60
Qy 61 CCGCGCGCGGACCTTGGCTGCTCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 CCGCGCGCGGACCTTGGCTGCTCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Qy 121 TCGAGCTGCG 180
Db 121 TCGAGCTGCG 180
Qy 181 CCGCGCGAGGAGCTGCGAGTACCTTGGGGGAGATCGAACHATCTCCGCTCGCGCGCTC 240
Db 181 CCGCGCGAGGAGCTGCGAGTACCTTGGGGGAGATCGAACHATCTCCGCTCGCGCGCTC 240
Qy 241 AGAAGAGCAGCCGACCGGACCGGACCGCTTGGGATTCGGGATTCGGGATTCGGGATTC 300
Db 241 AGAAGAGCAGCCGACCGGACCGGACCGCTTGGGATTCGGGATTCGGGATTCGGGATTC 300
Qy 301 AGCTCAGAGATGAAATCCCTCTGCTGACAAACAAAGCTGTCTGAAAGGCTGTCTTGG 360
Db 301 AGCTCAGAGATGAAATCCCTCTGCTGACAAACAAAGCTGTCTGAAAGGCTGTCTTGG 360
Qy 361 GAGGAGTCTCTGCTGTGTTATCAAGGATCCAAATGCTTAAAGAGCTGTCTTCAAGGGA 420
Db 361 GAGGAGTCTCTGCTGTGTTATCAAGGATCCAAATGCTTAAAGAGCTGTCTTCAAGGGA 420
Qy 421 GTCAAAATCTGGATGCGCAATGATCCCGAGACTTTTGGACAGCTGGGATTCACAC 480
Db 421 GTCAAAATCTGGATGCGCAATGATCCCGAGACTTTTGGACAGCTGGGATTCACAC 480
Qy 481 AGAAGAAAGGGGACTTGGGCGGAGTTATGCTTCCAGTGAGGCAATTTTGGGGCAGAA 540
Db 481 AGAAGAAAGGGGACTTGGGCGGAGTTATGCTTCCAGTGAGGCAATTTTGGGGCAGAA 540
Qy 541 TACAGAGATATGGAATCAGATTATCAGGACAGGAGTTGACCACTGCAAGAGTGATT 600
Db 541 TACAGAGATATGGAATCAGATTATCAGGACAGGAGTTGACCACTGCAAGAGTGATT 600
Qy 601 GACACATCAAAACCAACCTGACGACAGAGATCATCTGCTGCTTGGATTCACAAG 660
Db 601 GACACATCAAAACCAACCTGACGACAGAGATCATCTGCTGCTTGGATTCACAAG 660
Qy 661 GATCTCTCTGATGGCGCTCCATGCGCTCCATGCGCTCTGCGAGTCTATGTGGTGAAC 720
Db 661 GATCTCTCTGATGGCGCTCCATGCGCTCCATGCGCTCTGCGAGTCTATGTGGTGAAC 720
Qy 721 AGTGAGCTGTCTGCGAGCTGTACAGAGATCGGGAGACATGGGCTCGGTGGCTTTTC 780
Db 721 AGTGAGCTGTCTGCGAGCTGTACAGAGATCGGGAGACATGGGCTCGGTGGCTTTTC 780
Qy 781 AACATCGCAGCTACGCTCTCTCAGTACATGATTCGCAATCAGCGGCTCGAGCCA 840
Db 781 AACATCGCAGCTACGCTCTCTCAGTACATGATTCGCAATCAGCGGCTCGAGCCA 840
Qy 841 GGTGACTTATACACATTTGGGAGATGACATATTACTGAAATCATCGACCCACTG 900
Db 841 GGTGACTTATACACATTTGGGAGATGACATATTACTGAAATCATCGACCCACTG 900

Db 841 GGTGACTTATACACATTTGGGAGATGACATATTACTGAAATCATCGACCCACTG 900
Qy 901 AAAATTACGCTTCCAGGAGAACCCAGACCTTTCCCAAAGCTCAGGATTTCTTCCGAAAAGTT 960
Db 901 AAAATTACGCTTCCAGGAGAACCCAGACCTTTCCCAAAGCTCAGGATTTCTTCCGAAAAGTT 960
Qy 961 GAGAAATTTGATGATCTTCAAAGCTTGAAGACTTTTCAAGTTGAAGGCTCAATTCGCCATCCA 1020
Db 961 GAGAAATTTGATGATCTTCAAAGCTTGAAGACTTTTCAAGTTGAAGGCTCAATTCGCCATCCA 1020
Qy 1021 ACTATTAAAAATGAAAAATGGCTTTTAGGGTCTTTCAAAAGGAGCTTGAAGGATTTGCTCA 1080
Db 1021 ACTATTAAAAATGAAAAATGGCTTTTAGGGTCTTTCAAAAGGAGCTTGAAGGATTTGCTCA 1080
Qy 1081 GTCTTTAGGGGTTGGGCTTGGATGCGGAGGTAAAGTCTCTTTTGGCTCTTAAAGAAAGG 1140
Db 1081 GTCTTTAGGGGTTGGGCTTGGATGCGGAGGTAAAGTCTCTTTTGGCTCTTAAAGAAAGG 1140
Qy 1141 AACTAGTCAAAAAATCTCTCGTGACCTATCAGTTAATTTTAAAGGATTTGCGCACT 1200
Db 1141 AACTAGTCAAAAAATCTCTCGTGACCTATCAGTTAATTTTAAAGGATTTGCGCACT 1200
Qy 1201 GGCATAATGTAATGTCGCGAGTCTTTCCATAATAAAGGCTTTGAGTTAACTCACTGAGG 1260
Db 1201 GGCATAATGTAATGTCGCGAGTCTTTCCATAATAAAGGCTTTGAGTTAACTCACTGAGG 1260
Qy 1261 GTATCTGCAATGCTGAGGTATGACAAAGTGGAGAGTGAATGTATGTCTCTTAG 1320
Db 1261 GTATCTGCAATGCTGAGGTATGACAAAGTGGAGAGTGAATGTATGTCTCTTAG 1320
Qy 1321 CAAAAACATGATGTCATTTCAATCCACGTAATTTATAAAGAGGTTGGTGAATTTAC 1380
Db 1321 CAAAAACATGATGTCATTTCAATCCACGTAATTTATAAAGAGGTTGGTGAATTTAC 1380
Qy 1381 AAGCTATTTTGGAAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 1440
Db 1381 AAGCTATTTTGGAAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 1440
Qy 1441 CTGAGGAGCTGAGTAAACCATCGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1441 CTGAGGAGCTGAGTAAACCATCGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Qy 1501 TGTATTATATGTTGCTATAATAAAGAGTGTCTGC 1536
Db 1501 TGTATTATATGTTGCTATAATAAAGAGTGTCTGC 1536

RESULT 5
BC002567
LOCUS
DEFINITION
ACCSSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT

BC002567 1533 bp mRNA linear PRI 12-JUL-2001
Homo sapiens, thymidylate synthetase, clone MGC:1590 IMAGE:3138677,
mRNA, complete cds.
BC002567 GI:12803482
BC002567.1
MGC.
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1533)
Strausberg, R.
Direct Submission
Submitted (05-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgpps-remail.nih.gov
Tissue Procurement: ATCC
cdNA Library Preparation: Rubin Laboratory
cdNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural

QY 241 AGAAGAGCGACCGCAGCGGCACCGCACCCCTGTGCGTATTTCGGCATGCGAGCGCGCTAC 300
Db 241 AGAAGAGCGACCGCAGCGGCACCGCACCCCTGTGCGTATTTCGGCATGCGAGCGCGCTAC 300
QY 301 AGCCTGAGAGATGAATCCCTCTGCTGCAACCAACCGTGTGTTCTGGAAGGCTGTTTG 360
Db 301 AGCCTGAGAGATGAATCCCTCTGCTGCAACCAACCGTGTGTTCTGGAAGGCTGTTTG 360
QY 361 GAGGAGTTCGTGCTGTTTATCAAGGATCCCAAAATGCTAAAGAGCTGCTTCCAGGGA 420
Db 361 GAGGAGTTCGTGCTGTTTATCAAGGATCCCAAAATGCTAAAGAGCTGCTTCCAGGGA 420
QY 421 GTGAATCTGGATGCGCAATGATCCGAGACTTTTGGACAGCTGCGATTCCTCACC 480
Db 421 GTGAATCTGGATGCGCAATGATCCGAGACTTTTGGACAGCTGCGATTCCTCACC 480
QY 481 AGAAGAGAGGAGCTTGGGCCAGGTTTATGCTTCCAGTGGAGGCAATTTTGGGCGAGAA 540
Db 481 AGAAGAGAGGAGCTTGGGCCAGGTTTATGCTTCCAGTGGAGGCAATTTTGGGCGAGAA 540
QY 541 TACAGAGATATGGAATCAGATTATCAGACAGGAGTTCACCACTGCAAAAGAGTGATT 600
Db 541 TACAGAGATATGGAATCAGATTATCAGACAGGAGTTCACCACTGCAAAAGAGTGATT 600
QY 601 GACACCATCAAAACCAACCTGACGACAGAGATCATCATGTCGCTTGGATTCAGAA 660
Db 601 GACACCATCAAAACCAACCTGACGACAGAGATCATCATGTCGCTTGGATTCAGAA 660
QY 661 GATCTCTCTGATGCGCTGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 GATCTCTCTGATGCGCTGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 AGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 AGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 781 AACATGCGCAGCTACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 781 AACATGCGCAGCTACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 841 GGTGATTTATACACATTTGGGAGATGACATATTTACCTGATACATCATCAGGACCATG 900
Db 841 GGTGATTTATACACATTTGGGAGATGACATATTTACCTGATACATCATCAGGACCATG 900
QY 901 AAAATTCAGCTTACGAGAGAACCGACCTTCCAAAGCTCAGGATTTCTGAAAAAGTT 960
Db 901 AAAATTCAGCTTACGAGAGAACCGACCTTCCAAAGCTCAGGATTTCTGAAAAAGTT 960
QY 961 GAGAAATTCAGTCTCAAGCTGAGAGCTTTCAGATTCAGGCTACATCCGATCCA 1020
Db 961 GAGAAATTCAGTCTCAAGCTGAGAGCTTTCAGATTCAGGCTACATCCGATCCA 1020
QY 1021 ACTATTAATGGAATGCTGCTTGGGCTGCTTCAAGAGCTGCTTCAAGAGCTTGAAGGATTTGCA 1080
Db 1021 ACTATTAATGGAATGCTGCTTGGGCTGCTTCAAGAGCTGCTTCAAGAGCTTGAAGGATTTGCA 1080
QY 1081 GTCTTTAGGGCTTGGGCTGATGCCAGTAAAGTTCTTTTGTCTCTAAAGAAAGG 1140
Db 1081 GTCTTTAGGGCTTGGGCTGATGCCAGTAAAGTTCTTTTGTCTCTAAAGAAAGG 1140
QY 1141 AACTAGTCAAAATCTGTCGCTGACCTATCAGTTATTAATTTTTAAGGATGTTGCCACT 1200
Db 1141 AACTAGTCAAAATCTGTCGCTGACCTATCAGTTATTAATTTTTAAGGATGTTGCCACT 1200
QY 1201 GCGAAATGTAATGTCGCTGCTTTCATTAATAAAGGCTTTGAGTTAACTCACTGAGG 1260
Db 1201 GCGAAATGTAATGTCGCTGCTTTCATTAATAAAGGCTTTGAGTTAACTCACTGAGG 1260
QY 1261 GTATCTGCAATGCTGAGGTTATGACAAAGTGAAGGAGTAATGTTATGCTGCTCTTAG 1320
Db 1261 GTATCTGCAATGCTGAGGTTATGACAAAGTGAAGGAGTAATGTTATGCTGCTCTTAG 1320

QY 1321 CAAAACATGATGCTGCTTCAATCCCACTTCTTATAAAGAGGCTTGGTGAATTCAC 1380
Db 1321 CAAAACATGATGCTGCTTCAATCCCACTTCTTATAAAGAGGCTTGGTGAATTCAC 1380
QY 1381 AAGCTATTTTGGGAATATTTTAGAATATTTTAAAGATTTTCAAGCTATTCCTCAAAAT 1440
Db 1381 AAGCTATTTTGGGAATATTTTAGAATATTTTAAAGATTTTCAAGCTATTCCTCAAAAT 1440
QY 1441 CTGAGGAGCTGAGTAAACCATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1441 CTGAGGAGCTGAGTAAACCATCATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1501 TCTTTTATATGCTGCTATATAAAGAGGTTTCTGC 1536
Db 1501 TCTTTTATATGCTGCTATATAAAGAGGTTTCTGC 1536
RESULT 4
HSTS YN1
LOCUS HSTS YN1 1536 bp mRNA linear PRI 12-SEP-1993
DEFINITION Human mRNA for thymidylate synthase (EC 2.1.1.45).
ACCESSION X02308
VERSION X02308.1 GI:37478
KEYWORDS inverted repeat; synthetase; tandem repeat.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Theria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1536)
AUTHORS Takeishi, K., Kaneda, S., Ayusawa, D., Shimizu, K., Gotoh, O. and Seno, T.
TITLE Nucleotide sequence of a functional cDNA for human thymidylate synthase
JOURNAL Nucleic Acids Res. 13 (6), 2035-2043 (1985)
MEDLINE 85215597
PUBMED 2987839
COMMENT Data kindly reviewed (22-OCT-1985) by Seno T.
FEATURES
source
1. .1536
/location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature
14..103
/note="triple tandemly repeated elements"
repeat_region
14..16
/note="direct repeat 1"
misc_feature
35..103
/note="pot. stem-loop structure"
misc_feature
35..69
/note="pot. stem-loop structure"
repeat_unit
35..41
/note="inverted repeat A"
repeat_region
42..44
/note="direct repeat 1"
misc_feature
63..103
/note="pot. stem-loop structure"
repeat_unit
63..69
/note="inverted repeat A"
repeat_region
70..72
/note="direct repeat 1"
repeat_unit
97..103
/note="inverted repeat A'"
repeat_region
104..106
/note="direct repeat 1"
106..1047
/note="thymidylate synthase (aa 1-313)"
/codon_start=1
/protein_id="CAA26178.1"
/db_xref="GI:37479"
/db_xref="SWISS-PROT:P04818"
/translation="MPVAGSELPRPLPPAAQERDAEPPPHGELQYLQIQHILRCG
VRKDDRTGTLISVFGMQARYSLRDEPPLTTTKRVFKGVLEELLWFLKSTNAKELS
SKGVKTDANGSRDFLSLGFTREGLGVYQFWRHFGAEYRDMESYSGQVQDQ
LQVLDITKTNPDERRITWCANPNRDLPLMALPPCHALCQFYVNVNSELSCQLYQRSD

Sun Dec 1 18:15:45 2002

/tissue type="Bone marrow, chronic myelogenous leukemia"
/clone lib="NIH MGC_54"
/lab_host="DH10B"
/notes="Vector: pDNR-LIB"
96...1037
/codon_start=1
/product="similar to thymidylate synthetase"
/protein_id="AAH13919.1"
/db_xref="GI:15530267"
/translation="MPVAGSELPRPPLPPAQAERDAERPPHGELOVLQIHLRTG
VRKDRITGTLVSFQMAQYSLRDEPLLTKRVFWKLELLNFIKSTNAKES
SKGKIWDANGSRDLGLSFTREBDLGPVQFWRHFGABYRDMESYSGQVQ
LQVIDITKINPDNRRRIIMCAWNPDLPLPCHALCOFVNVNSLSQLYQRSGD
MGLGVFPFIASVALLTYMIAHTLGLKPGDFIHLGLDAHYLNHLEPLKIQLOREPRPF
PKRLILRKVEKIDDFKAEQDFQEGYNPHPTIKEMAV"

CDS	425 a	368 c	398 g	378 t
BASE COUNT	425 a	368 c	398 g	378 t
ORIGIN	88.4%; Score 1358; DB 9; Length 1569;			
Query Match	88.4%; Score 1358; DB 9; Length 1569;			
Best Local Similarity	99.9%; Pred. No. 0;			
Matches 1458; Conservative	0; Mismatches 2; Indels 0; Gaps 0;			
QY	37	CGCGCCACTTGGCCTCGCTCCGCGCGCCACTTCGCTCGCTCGCTCCCGCCGCC 96		
DB	27	CGCGCCACTTGGCCTCGCTCCGCGCGCCACTTCGCTCGCTCGCTCCCGCCGCC 86		
QY	97	CGCGCGCCACTTGGCCTCGCTCCGCGCGCCACTTCGCTCGCTCGCTCCCGCCGCC 156		
DB	87	CGCGCGCCACTTGGCCTCGCTCCGCGCGCCACTTCGCTCGCTCGCTCCCGCCGCC 146		
QY	157	CAGGAGCGGAGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 216		
DB	147	CAGGAGCGGAGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 206		
QY	217	CAACACATCTCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 276		
DB	207	CAACACATCTCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 266		
QY	277	GTATTCGGCATGACGCGCGCTACAGCTGAGAGATGAGATGAGATGAGATGAGAT 336		
DB	267	GTATTCGGCATGACGCGCGCTACAGCTGAGAGATGAGATGAGATGAGATGAGAT 326		
QY	337	CGTGTCTTCTGAGAGGTGTTTGGAGAGGTGTTTGGAGAGGTGTTTGGAGAGGT 396		
DB	327	CGTGTCTTCTGAGAGGTGTTTGGAGAGGTGTTTGGAGAGGTGTTTGGAGAGGT 386		
QY	397	GCTAAAGAGGTGTTTCCAAAGGAGTGAATCTGGAGTGAATCTGGAGTGAATCT 456		
DB	387	GCTAAAGAGGTGTTTCCAAAGGAGTGAATCTGGAGTGAATCTGGAGTGAATCT 446		
QY	457	TTGGACAGCTGGGATTTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 516		
DB	447	TTGGACAGCTGGGATTTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 506		
QY	517	CAGTGGAGGATTTTGGGGGAGAGATACAGAGATATGGAATCAGATATGGAATC 576		
DB	507	CAGTGGAGGATTTTGGGGGAGAGATACAGAGATATGGAATCAGATATGGAATC 566		
QY	577	GTTGACCAACTGCAAGAGTATTTGACCACTCAAAACCAACCTGACGACAGAGATC 636		
DB	567	GTTGACCAACTGCAAGAGTATTTGACCACTCAAAACCAACCTGACGACAGAGATC 626		
QY	637	ATCATGTGGCTTGAATTCGAAGAGATTTCTCTGATGGCGCTGCTCCATGCCATGCC 696		
DB	627	ATCATGTGGCTTGAATTCGAAGAGATTTCTCTGATGGCGCTGCTCCATGCCATGCC 686		
QY	697	CTCTGCGAGTCTATGTGTGAACAGTGAAGTGTCTGCGAGCTGTCTGCGAGCTGT 756		
DB	687	CTCTGCGAGTCTATGTGTGAACAGTGAAGTGTCTGCGAGCTGTCTGCGAGCTGT 746		
QY	757	GACATGGGCTCGGTGTGCTTTTCAACATCGCGAGTGTGCGGCTGCTCAGCTACGAT 816		
DB	747	GACATGGGCTCGGTGTGCTTTTCAACATCGCGAGTGTGCGGCTGCTCAGCTACGAT 806		

QY	817	GGGCACATCAGCGGCTCGAGCCAGGTGACTTTATACACATTTGGGAGTGCATATT 876		
DB	807	GGGCACATCAGCGGCTCGAGCCAGGTGACTTTATACACATTTGGGAGTGCATATT 866		
QY	877	TACTGTAATCATCTGAGGCTGAAATTCAGCTTCAGGAGACCCAGACCTTTCCCA 936		
DB	867	TACTGTAATCATCTGAGGCTGAAATTCAGCTTCAGGAGACCCAGACCTTTCCCA 926		
QY	937	TAAGCTCAGGATTTCTGCAAAAGTTCAGAAATTCAGCTTCAGGAGTGAAGCTTTTCAG 996		
DB	927	TAAGCTCAGGATTTCTGCAAAAGTTCAGAAATTCAGCTTCAGGAGTGAAGCTTTTCAG 986		
QY	997	ATTGAGGATTAATCTCGGATCCACTTAAATTAATTAATTAATTAATTAATTAATTA 1056		
DB	987	ATTGAGGATTAATCTCGGATCCACTTAAATTAATTAATTAATTAATTAATTAATTA 1046		
QY	1057	TAAGGAGCTGAGGATTTGTCAGCTCTTTAGGGGTTGGGCTGGATGCCGAGTAAAGT 1116		
DB	1047	TAAGGAGCTGAGGATTTGTCAGCTCTTTAGGGGTTGGGCTGGATGCCGAGTAAAGT 1106		
QY	1117	TCTTTTCTCTTAAAGAAAGAAAGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1176		
DB	1107	TCTTTTCTCTTAAAGAAAGAAAGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1166		
QY	1177	TCTTTTCTCTTAAAGAAAGAAAGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1236		
DB	1167	TCTTTTCTCTTAAAGAAAGAAAGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1226		
QY	1237	AGGCTTGTAGTTAACTCACTGAGGATCTCAATCTGCTGCTGCTGCTGCTGCTGCTGCT 1296		
DB	1227	AGGCTTGTAGTTAACTCACTGAGGATCTCAATCTGCTGCTGCTGCTGCTGCTGCTGCT 1286		
QY	1297	AGGCTTGTAGTTAACTCACTGAGGATCTCAATCTGCTGCTGCTGCTGCTGCTGCTGCT 1356		
DB	1287	AGGCTTGTAGTTAACTCACTGAGGATCTCAATCTGCTGCTGCTGCTGCTGCTGCTGCT 1346		
QY	1357	ATTAAGAGGCTGAGGATTTTCAAGCTATTTTGGATATTTTGGATATTTTGGATATTTT 1416		
DB	1347	ATTAAGAGGCTGAGGATTTTCAAGCTATTTTGGATATTTTGGATATTTTGGATATTTT 1406		
QY	1417	ATTTCACAGCTATTTCCCTCAAACTGAGGAGCTGAGTAAACCATCATCATGATGTA 1476		
DB	1407	ATTTCACAGCTATTTCCCTCAAACTGAGGAGCTGAGTAAACCATCATCATGATGTA 1466		
QY	1477	GAGTGTGGTTATGAACTTTA 1496		
DB	1467	GAGTGTGGTTATGAACTTTA 1486		

RESULT 6
BC002567 1533 bp mRNA linear PRI 12-JUL-2002
LOCUS Homo sapiens, thymidylate synthetase, clone MGC:1590 IMAGE:31388
DEFINITION mRNA, complete cds.
ACCESSION BC002567
VERSION BC002567.1 GI:12803482
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1533)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

Best Local Similarity 99.9%; Pred. No. 0;
Matches 1494; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1	GGGGGGGGGGGACACACTTGGCTTGCCTCCGTCGCGCGCGCACATTGGCTTGCCTCCGT	60
1	GGGGGGGGGGGGGACACACTTGGCTTGCCTCCGTCGCGCGCGCACATTGGCTTGCCTCCGT	60
61	CCGCGCGCGCCACTTTCGCGCTGCCTCCGTCGCGCGCGCGCGCCATGCTGTGGCGCGC	120
61	CCGCGCGCGCGCACTTTCGCGCTGCCTCCGTCGCGCGCGCGCGCATGCTGTGGCGCGC	120
121	TGGAGCTGCGCGCGCGCGCTTTCGCGCGCGCGCGCGCGCAAGAGAGGGAGCGCGAGCCGCT	180
121	TGGAGCTGCGCGCGCGCGCGCTTTCGCGCGCGCGCGCGCGCAAGAGAGGGAGCGCGAGCCGCT	180
181	CCGCGCGAGCGGGAGCTGCAGTACCTTGGGCGAGATCCAAACACATCCTCGCTGGGCGCTC	240
181	CCGCGCGAGCGGGAGCTGCAGTACCTTGGGCGAGATCCAAACACATCCTCGCTGGGCGCTC	240
241	AGGAGAGAGCGACGCGACCGGCGACCGGCACCTGTGCGTATTTCGGCATTCAGGCGCGCTAC	300
241	AGGAGAGAGCGACGCGACCGGCGACCGGCACCTGTGCGTATTTCGGCATTCAGGCGCGCTAC	300
301	AGCCTGAGAGATGAATTCCTCTGCTGACAAACAAACGCTGTGTTGGAAGGCTGTTTG	360
301	AGCCTGAGAGATGAATTCCTCTGCTGACAAACAAACGCTGTGTTGGAAGGCTGTTTG	360
361	GAGGAGTGTGCTGTGTTTATCAAGGATCCAAATGCTTAAAGCTGCTTCCAAAGGA	420
361	GAGGAGTGTGCTGTGTTTATCAAGGATCCAAATGCTTAAAGCTGCTTCCAAAGGA	420
421	GTGAAATCTGGGATCCAAATGATCCCGAGACTTTTGGACAGCTTGGGATTCCTCAAC	480
421	GTGAAATCTGGGATCCAAATGATCCCGAGACTTTTGGACAGCTTGGGATTCCTCAAC	480
481	AGAGAGAGAGGAGCTTGGGCGCGAGTTATGCGCTTCCAGTGGAGGCAATTTGGGCGAGAA	540
481	AGAGAGAGAGGAGCTTGGGCGCGAGTTATGCGCTTCCAGTGGAGGCAATTTGGGCGAGAA	540
541	TACAGAGATATGGAATCAGATTATTTCAGGACAGGAGTTGACCACTGCGAAAGAGTGATT	600
541	TACAGAGATATGGAATCAGATTATTTCAGGACAGGAGTTGACCACTGCGAAAGAGTGATT	600
601	GACACATCAAAACCAACCTCAACAGACAGGAATCATCATGTCGCTGGGAATCCAAGA	660
601	GACACATCAAAACCAACCTCAACAGACAGGAATCATCATGTCGCTGGGAATCCAAGA	660
661	GATCTTCTCTGATGCGGTGCTCCATGCGATGCGCTTCCAGTTCATGTGTGAAC	720
661	GATCTTCTCTGATGCGGTGCTCCATGCGATGCGCTTCCAGTTCATGTGTGAAC	720
721	AGTGAGCTGTCTGCCAGCTGTACCAGAGATCGGAGACATGGGCTCGGTGTGCTTTC	780
721	AGTGAGCTGTCTGCCAGCTGTACCAGAGATCGGAGACATGGGCTCGGTGTGCTTTC	780
781	AACATCGCGAGCTACGCGCTGCTACATGATTTGCGCATCACGCGGCGCTGAAGCA	840
781	AACATCGCGAGCTACGCGCTGCTACATGATTTGCGCATCACGCGGCGCTGAAGCA	840
841	GGTGACTTTATACACACTTTGGGAGATGACATTTTACCTGAATCACATCGAGCACTG	900
841	GGTGACTTTATACACACTTTGGGAGATGACATTTTACCTGAATCACATCGAGCACTG	900
901	AAATTCAGCTTACGAGAGACCGAGCTTTCCTCCAAAGCTCAGGATTTCTCGAAAAGTT	960
901	AAATTCAGCTTACGAGAGACCGAGCTTTCCTCCAAAGCTCAGGATTTCTCGAAAAGTT	960
961	GAGAAAATTGATGACTTCAAAGCTGAGAGCTTTTCAAGAGGCTGCAATTCGGATCCA	1020
961	GAGAAAATTGATGACTTCAAAGCTGAGAGCTTTTCAAGAGGCTGCAATTCGGATCCA	1020
1021	ACTATTAATTAAGAAATGCGCTTTTGGGTGCTTTTCAAGAGGCTGCAATTCGGATCCA	1080

1021	dbb	ACTATTAAATGGAATGGCTGTGGTGGCTTTCAAAGGAGCTTTGGGATATTGTCA	1080
1081	QOY	GTCTTTTAGGGGTTGGCTGGATCGCGAGGTAAAGTCTCTTTTGTCTCTAAAAGAAAGG	1140
1081	dbb	GTCTTTTAGGGGTTGGCTGGATCGCGAGGTAAAGTCTCTTTTGTCTCTAAAAGAAAGG	1140
1141	QOY	AACTAGGTCAAAAATCTGTCGGTGACCTATCATGTTATTAAATTTTAAAGGATGTTGCCACT	1200
1141	dbb	AACTAGGTCAAAAATCTGTCGGTGACCTATCATGTTATTAAATTTTAAAGGATGTTGCCACT	1200
1201	QOY	GGCAAAATGAACGTGCCAGTTCCTTCCATAATAAAGGCTTTGAGTTAACTCACTGAGG	1260
1201	dbb	GGCAAAATGAACGTGCCAGTTCCTTCCATAATAAAGGCTTTGAGTTAACTCACTGAGG	1260
1261	QOY	GTATCTGCAAACTGCTGAGGTATTGAACAAAGTGAGGAGAAATGAATGTATGCTCTTAG	1320
1261	dbb	GTATCTGCAAACTGCTGAGGTATTGAACAAAGTGAGGAGAAATGAATGTATGCTCTTAG	1320
1321	QOY	CAAAAACATGTATGTCCATTTCAATCCACGTACTATTAAAGAGGTTGGTGAATTTTCA	1380
1321	dbb	CAAAAACATGTATGTCCATTTCAATCCACGTACTATTAAAGAGGTTGGTGAATTTTCA	1380
1381	QOY	AAGCTATTTTGGAAATATTTTGAATATTTTAAAGATTTTCAAGCTATTTCCCTCAAAAT	1440
1381	dbb	AAGCTATTTTGGAAATATTTTGAATATTTTAAAGATTTTCAAGCTATTTCCCTCAAAAT	1440
1441	QOY	CTGAGGAGAGCTGAGTAAACCAATCGATCATGATGTAGAGTGTGGTTATGAACATTTA	1496
1441	dbb	CTGAGGAGAGCTGAGTAAACCAATCGATCATGATGTAGAGTGTGGTTATGAACATTTA	1496

RESULT 5				
BC013919				
LOCUS				
DEFINITION	BC013919	1569 bp	mRNA	linear
	Homo sapiens,			PRI 10-SEP-
	IMAGE:4048625,			clone MGC:2288-
	mRNA, complete cds.			

ACCESSION	BC013919	
VERSION	BC013919.1	GI:15530266
KEYWORDS	MGC.	
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa;	Cnidata; Vertebrata; Euteleostom
	Mammalia; Eutheria;	Catarrhini; Hominoideae; Homo.
REFERENCE	1 (bases 1 to 1569)	

REFERENCE
1 (bases 1 to 1569)
Strausberg R.
AUTHORS
TITLE
Direct Submission
JOURNAL
Submitted (07-SEP-2001) National Institutes of Health, Mammalia
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590
USA

REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact : MGC help desk Email: csapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: CLONTECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbiology.org contact: amadan@systemsbiology.org Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Juli Greene, Mark Kettman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRAL Plate: 31 Row: 0 Column: 4
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

```

FEATURES
  source
    1. 1569
      Location/Qualifiers
        to_protein:
          tag:
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="MGC:22884 IMAGE:4048625"

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

from nucleic - nucleic search, using sw model

Run on: November 28, 2002, 19:48:24 ; Search time 119.586 Seconds
(without alignments)
8693.659 Million cell updates/sec

Title: US-09-963-333-6
Perfect score: 1536
Sequence: 1 ggggggggggggaccacttg.....ataataagaagtgttctgc 1536

Scoring table: IDENTITY_NUC
Gapop 10.0 : Gapext 1.0

Searched: 344316 seqs; 338423730 residues

total number of hits satisfying chosen parameters: 688632

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```

Database : Pending_Patents_NA_New.*
1: /cgn2_6/ptodata/2/pna/pct_NEW_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is defined by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1533	99.8	1536	5	US-09-658-659B-200		Sequence 6, Appli
2	1493.4	97.2	3298	6	US-10-240-965-200		Sequence 200, App
C	1126.4	73.3	1942	5	US-09-724-676-34523		Sequence 34523, A
	1126.4	73.3	1942	5	US-09-724-676A-34523		Sequence 34523, A
C	1126.4	73.3	2042	5	US-09-724-676-34515		Sequence 34515, A
C	1126.4	73.3	2042	5	US-09-724-676A-34515		Sequence 34515, A
C	1126.4	73.3	2042	5	US-09-724-676-34597		Sequence 34597, A
C	1126.4	73.3	2391	5	US-09-724-676A-34597		Sequence 34597, A
C	1126.4	73.3	2391	5	US-09-724-676-34535		Sequence 34535, A
C	1126.4	73.3	2411	5	US-09-724-676A-34535		Sequence 34535, A
C	1126.4	73.3	2411	5	US-09-724-676-34589		Sequence 34589, A
C	1126.4	73.3	2491	5	US-09-724-676A-34589		Sequence 34589, A
C	1126.4	73.3	2527	5	US-09-724-676-34548		Sequence 34548, A
C	1126.4	73.3	2527	5	US-09-724-676A-34548		Sequence 34548, A
C	1126.4	73.3	3023	5	US-09-724-676-34565		Sequence 34565, A
C	1126.4	73.3	3023	5	US-09-724-676A-34565		Sequence 34565, A
C	1126.4	73.3	3123	5	US-09-724-676-34557		Sequence 34557, A
C	1126.4	73.3	3123	5	US-09-724-676A-34557		Sequence 34557, A
C	1126.4	73.3	3139	5	US-09-724-676-34581		Sequence 34581, A
C	1126.4	73.3	3139	5	US-09-724-676A-34581		Sequence 34581, A
C	1126.4	73.3	3239	5	US-09-724-676-34573		Sequence 34573, A
C	1126.4	73.3	3239	5	US-09-724-676A-34573		Sequence 34573, A
C	1126.4	73.3	3239	5	US-09-724-676-34524		Sequence 34524, A
C	928.8	60.5	2083	5	US-09-724-676A-34516		Sequence 34516, A
C	928.8	60.5	2183	5	US-09-724-676-34516		Sequence 34516, A
C	928.8	60.5	2183	5	US-09-724-676A-34516		Sequence 34516, A

C	27	928.8	60.5	2552	5	US-09-724-676-34536	Sequence 34536, A
C	28	928.8	60.5	2552	5	US-09-724-676A-34536	Sequence 34536, A
C	29	928.8	60.5	2632	5	US-09-724-676-34590	Sequence 34590, A
C	30	928.8	60.5	2632	5	US-09-724-676A-34590	Sequence 34590, A
C	31	928.8	60.5	2652	5	US-09-724-676-34529	Sequence 34529, A
C	32	928.8	60.5	2652	5	US-09-724-676A-34529	Sequence 34529, A
C	33	928.8	60.5	2668	5	US-09-724-676-34549	Sequence 34549, A
C	34	928.8	60.5	2668	5	US-09-724-676A-34549	Sequence 34549, A
C	35	928.8	60.5	2768	5	US-09-724-676-34542	Sequence 34542, A
C	36	928.8	60.5	2768	5	US-09-724-676A-34542	Sequence 34542, A
C	37	928.8	60.5	3164	5	US-09-724-676-34566	Sequence 34566, A
C	38	928.8	60.5	3164	5	US-09-724-676A-34566	Sequence 34566, A
C	39	928.8	60.5	3253	5	US-09-724-676-34525	Sequence 34525, A
C	40	928.8	60.5	3253	5	US-09-724-676A-34525	Sequence 34525, A
C	41	928.8	60.5	3264	5	US-09-724-676-34558	Sequence 34558, A
C	42	928.8	60.5	3264	5	US-09-724-676A-34558	Sequence 34558, A
C	43	928.8	60.5	3280	5	US-09-724-676-34582	Sequence 34582, A
C	44	928.8	60.5	3280	5	US-09-724-676A-34582	Sequence 34582, A
C	45	928.8	60.5	3353	5	US-09-724-676-34517	Sequence 34517, A

ALIGNMENTS

```

, RESULT 1
, US-09-658-659B-6
, Sequence 6, Application US/09658659B
, GENERAL INFORMATION:
, APPLICANT: Stanton, Jr., Vincent P.
, TITLE OF INVENTION: GENE SEQUENCE VARIANCE IN GENES RELATED
, TITLE OF INVENTION: TO FOLATE METABOLISM HAVING UTILITY IN DETERMINING THE
, TITLE OF INVENTION: TREATMENT OF DISEASE
, FILE REFERENCE: 11926-015001
, CURRENT APPLICATION NUMBER: US/09/658,659B
, CURRENT FILING DATE: 2000-09-08
, PRIOR APPLICATION NUMBER: 09/596,033
, PRIOR FILING DATE: 2000-06-15
, PRIOR APPLICATION NUMBER: 09/357,743
, PRIOR FILING DATE: 1999-07-20
, PRIOR APPLICATION NUMBER: 09/357,024
, PRIOR FILING DATE: 1999-07-19
, PRIOR APPLICATION NUMBER: 60/093,484
, PRIOR FILING DATE: 1998-07-20
, NUMBER OF SEQ ID NOS: 16
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 6
, LENGTH: 1536

```

```

; LENGTH: 1338
; TYPE: DNA

```

```

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1066
OTHER INFORMATION: n = t or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1136
OTHER INFORMATION: n = a or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1497
OTHER INFORMATION: n = t or a
US-09-6558-6598-6

```

Source: Match

Query Match	99.8%;	Score 1333;	DB 3;	Length 1336;
Best Local Similarity	100.0%;	Pred. No. 0;		
Mismatches 1536.	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

1 GGGGGGGGGGACCACTTGCCCTGCCCTCCGCCGCCACTTGGCCTGCCCTCCGT 60

1 aaccccccgcgcgaattggcctgctccgccgccacattggcctgacctcgt 60

Qy 61 CCGCGCGGCCAATTGGCTTGGCTCCGTCCCGCCCGCGCGGCCATGCCTGTGGCCGSC 120

Db 121 TTGCCCCCGCCGACAGAGGCGGAGCGCGCGCGTCCGCGCGCACGGGAGCTGCAG 180
Qy 202 TACCTGGGCGAGATCCAAACATCTCCGCTCGCGCTCAGGAGGACGACCGCACGGGC 261
Db 181 TACCTGGGCGAGATCCAAACATCTCCGCTCGCGCTCAGGAGGACGACCGCACGGGC 240
Qy 262 ACCGCGACCCCTGTGCTGATTCGGCATCGAGGCGCGCTACAGCCTGAGAGATGAATCCCT 321
Db 241 ACCGCGACCCCTGTGCTGATTCGGCATCGAGGCGCGCTACAGCCTGAGAGATGAATCCCT 300
Qy 322 CTGCTGACACAAACGCTGTTCTCGAAGGCTGTTTGGAGGAGTGTCTGTTATC 381
Db 301 CTGCTGACACAAACGCTGTTCTCGAAGGCTGTTTGGAGGAGTGTCTGTTATC 360
Qy 382 AAGGATCCAAATGCTAAAGAGCTGTCTTCAAGGAGTGAATTCGGATGCCAT 441
Db 361 AAGGATCCAAATGCTAAAGAGCTGTCTTCAAGGAGTGAATTCGGATGCCAT 420
Qy 442 GGATCCGAGACTTTTGGACAGCTGGATCTCCACGAGAGAGAGGAGCTTGGC 501
Db 421 GGATCCGAGACTTTTGGACAGCTGGATCTCCACGAGAGAGAGGAGCTTGGC 480
Qy 502 CAGTTTATGCTTCCAGTGGGSCATTTTGGGCGAGATACAGAGATATGAATCAGAT 561
Db 481 CAGTTTATGCTTCCAGTGGGSCATTTTGGGCGAGATACAGAGATATGAATCAGAT 540
Qy 562 TATTCAGGACAGGAGTTCACCACTGCAAGAGTATTCACCATCAAAACCAACCT 621
Db 541 TATTCAGGACAGGAGTTCACCACTGCAAGAGTATTCACCATCAAAACCAACCT 600
Qy 622 GAGCAGAGAGATCATCATGTGGCTTGAATCCAGAGATCTTCTCTGATGGCGTG 681
Db 601 GAGCAGAGAGATCATCATGTGGCTTGAATCCAGAGATCTTCTCTGATGGCGTG 660
Qy 682 CTCTCATGCGCATCTGCTGCTGCTATGCTGAGAGAGAGTGTCTCTGAGAGTG 741
Db 661 CTCTCATGCGCATCTGCTGCTGCTATGCTGAGAGAGAGTGTCTCTGAGAGTG 720
Qy 742 TACCAGAGATCCGAGACATGGGCTCGTGTGCTTCAACATCGCGAGCTAGCCCTG 801
Db 721 TACCAGAGATCCGAGACATGGGCTCGTGTGCTTCAACATCGCGAGCTAGCCCTG 780
Qy 802 CTCAGTATCATATTCGCGCATCAACGCGCTCAGCGAGTGTCTTATACACACTTG 861
Db 781 CTCAGTATCATATTCGCGCATCAACGCGCTCAGCGAGTGTCTTATACACACTTG 840
Qy 862 GGAGTGCATATTTACTGTAATCATCGAGCTCAAAATTCAGCTTCAGCGAGAA 921
Db 841 GGAGTGCATATTTACTGTAATCATCGAGCTCAAAATTCAGCTTCAGCGAGAA 900
Qy 922 CCGAGACTTTCGCAAGCTCAGGATCTTCGAAAAGTTGAGAAAATTCAGCTTCAA 981
Db 901 CCGAGACTTTCGCAAGCTCAGGATCTTCGAAAAGTTGAGAAAATTCAGCTTCAA 960
Qy 982 GCTGAGACTTTCAGATTGAAGGTTACATCCGATCCAAATTAATTAATGGAATGCT 1041
Db 961 GCTGAGACTTTCAGATTGAAGGTTACATCCGATCCAAATTAATTAATGGAATGCT 1020
Qy 1042 GTTTPAGGCTTTCAGAGGCTGAGATATTCAGTCTTTAGGGTGGCTGGA 1101
Db 1021 GTTTPAGGCTTTCAGAGGCTGAGATATTCAGTCTTTAGGGTGGCTGGA 1080
Qy 1102 TGCGAGGTAAGATCTTTTCTTAAAGANAGGAATAGTCAAAAATCTGTCC 1161
Db 1081 TGCGAGGTAAGATCTTTTCTTAAAGANAGGAATAGTCAAAAATCTGTCC 1140
Qy 1162 GTGACCTATCAGTTATTAATTTTAAAGATCTTCCACTGCGAAATGAACTGTGCCAGT 1221
Db 1141 GTGACCTATCAGTTATTAATTTTAAAGATCTTCCACTGCGAAATGAACTGTGCCAGT 1200
Qy 1222 TCTTTCCATATAAAGGCTTTGAGTTAACTCACTGAGGGTATCTGACATCTGAGTT 1281
Db 1201 TCTTTCCATATAAAGGCTTTGAGTTAACTCACTGAGGGTATCTGACATCTGAGTT 1260

Qy 1282 ATGACAAAGTGGAGAGATCAATATGATGTCTTTAGCAAAACATGATGTGCAATTT 1341
Db 1261 ATGACAAAGTGGAGAGATCAATATGATGTCTTTAGCAAAACATGATGTGCAATTT 1320
Qy 1342 CAATCCACGTAATTTATAAAGAGTGGTGAATTTCAAGCTATTTTGGAAATATTT 1401
Db 1321 CAATCCACGTAATTTATAAAGAGTGGTGAATTTCAAGCTATTTTGGAAATATTT 1380
Qy 1402 TAGAATATTTTAAAGATTTCAAGCTATTTCCCTCAAAATCTGAGGAGCTGAGTAACACC 1461
Db 1381 TAGAATATTTTAAAGATTTCAAGCTATTTCCCTCAAAATCTGAGGAGCTGAGTAACACC 1440
Qy 1462 ATCGATCATGATGTAGAGTGGTATGAATCTTAAAGATTAAGTTTATATGATGTGC 1515
Db 1441 ATCGATCATGATGTAGAGTGGTATGAATCTTAAAGATTAAGTTTATATGATGTGC 1500
Qy 1516 TATAATAAAGAGTGTCTGC 1536
Db 1501 TATAATAAAGAGTGTCTGC 1521

RESULT 3
US-09-724-676-34523/c
; Sequence 34523, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 34523
; LENGTH: 1942
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-34523

Query Match 73.3%; Score 1126.4; DB 5; Length 1942;
Best Local Similarity 98.0%; Pred. No. 2.6e-275;
Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGACCATGTCCTCCGTCGCGCGGCGCATCTGGCTGCTCCGT 60
Db 1942 GGGGGGGGGGGGACCATGTCCTCCGTCGCGCGGCGCATCTGGCTGCTCCGT 1883
Qy 61 CCGCGCGCGCGCATCTGGCTGCTCCGTCGCGCGGCGCGCATCTGGCTGCGCGGC 120
Db 1882 CCGCGCGCGCGCATCTGGCTGCTCCGTCGCGCGGCGCGCATCTGGCTGCGCGGC 1823
Qy 121 TCGAGCTGCGCGCGCGCATCTGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCT 180
Db 1822 TCGAGCTGCGCGCGCGCGCATCTGGCTGCGCGCGCGCGCGCGCGCGCGCGCT 1763
Qy 181 CCGCGCGCGCGCGCGCATCTGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCTC 240
Db 1762 CCGCGCGCGCGCGCGCATCTGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCTC 1703
Qy 241 AGGAAGACCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTAC 300
Db 1702 AGGAAGACCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTAC 1643
Qy 301 AGCTGAGAGATGAATTCCTCTGCTGACAAACAAAGCTGTCTTGAAGGGTGTTCG 360
Db 1642 AGCTGAGAGATGAATTCCTCTGCTGACAAACAAAGCTGTCTTGAAGGGTGTTCG 1583
Qy 361 GAGGAGTGTCTGTTTATCAAGGATCCAAATTCCTAAAGAGCTCTCTTCCAGGGA 420
Db 1582 GAGGAGTGTCTGTTTATCAAGGATCCAAATTCCTAAAGAGCTCTCTTCCAGGGA 1523
Qy 421 GTGAAATCTGGATGCCAATGGATCCGAGCTTTTGGACAGCTTGGATCTCCACC 480

```
Db 1522 GTCAAAATCTGGGATGCCAATSGATCCGAGACTTTTGGACAGCCTGGGATTTCTCCACC 1463
QY 481 AGAGAGAGAGGAGACTTGGGCCCGAGTTTATGGCTTCCAGTGGAGCAATTTGGGGAGAA 540
Db 1462 AGAGAGAGAGGAGACTTGGGCCCGAGTTTATGGCTTCCAGTGGAGCAATTTGGGGAGAA 1403
QY 541 TACAGAGATATGAATCAGATTATTCAGGACAGGAGTTCACCACTGCAAAAGAGTGATT 600
Db 1402 TACAGAGATATGAATCAGATTATTCAGGACAGGAGTTCACCACTGCAAAAGAGTGATT 1343
QY 601 GACACCATCAAAACCAACCCCTGACGACAGAGAAATCATCATGTGGCTTGGATTCGAAGA 660
Db 1342 GACACCATCAAAACCAACCCCTGACGACAGAGAAATCATCATGTGGCTTGGATTCGAAGA 1283
QY 661 GATCTTCTCTGATGGGCTGCTTCCATGCGATGCGCTTCCAGCTTCTATGTGGTGAAC 720
Db 1282 GATCTTCTCTGATGGGCTGCTTCCATGCGATGCGCTTCCAGCTTCTATGTGGTGAAC 1223
QY 721 AGTGAGCTGTCTGACACTGTACAGAGATCGGAGACATCGGCTCGGTGCTTTTC 780
Db 1222 AGTGAGCTGTCTGACACTGTACAGAGATCGGAGACATCGGCTCGGTGCTTTTC 1163
QY 781 AACATCGCCAGCTACGCCCTCTCACTGATCATGATTGGCAGATCAACCGGCTGAAGCCA 840
Db 1162 AACATCGCCAGCTACGCCCTCTCACTGATCATGATTGGCAGATCAACCGGCTGAAGCCA 1103
QY 841 GGTGACTTTATACACATTTTGGGAGATGACATATTTACTTGAATCATCATCGACCACTG 900
Db 1102 GGTGACTTTATACACATTTTGGGAGATGACATATTTACTTGAATCATCATCGACCACTG 1043
QY 901 AAAATTCAGCTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGGATTTCTGAAAAGTT 960
Db 1042 AAAATTCAGCTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGGATTTCTGAAAAGTT 983
QY 961 GAGAAAATGATGACTTCAAGCTGAAGACTTTCAGATTTCAAGGGTACAAATCCGCATCCA 1020
Db 982 GAGAAAATGATGACTTCAAGCTGAAGACTTTCAGATTTCAAGGGTACAAATCCGCATCCA 923
QY 1021 ACTATTAATGAAATGGCTTTTGGGCTGCTTTTAAAGAGCTGGAAGATTTGTCA 1080
Db 922 ACTATTAATGAAATGGCTTTTGGGCTGCTTTTAAAGAGCTGGAAGATTTGTCA 863
QY 1081 GTCTTTAGGGGTTGGGCTGGAATGCCAGGATTAAGATTTCTTTTGTCTTAAAGANAAG 1140
QY 1141 AACTAGTCAAAAATCTCTCCG 1162
Db 802 AGGGTGTATGACTTCTAGCTG 781

RESULT 4
US-09-724-676A-34523/c
; Sequence 34523, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34523
; LENGTH: 1942
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-34523

Query Match 73.3%; Score 1126.4; DB 5; Length 1942;
Best Local Similarity 98.0%; Pred. No. 2.6e-275;
Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY 1 GGGGGGGGGGGGACCACTTGGCTGCTCGTCCGCGCGGCCCACTTGGCTGCTCGCT 60
```

```
Db 1942 GGGGGGGGGGGGACCACTTGGCTGCTCGTCCGCGCGGCCCACTTGGCTGCTCGCT 1883
QY 61 CCGCGCGCGCGCACTTGGCTGCTCGTCCGCGCGGCCCGCGCGCATGCTGTGGCGCGC 120
Db 1882 CCGCGCGCGCGCACTTGGCTGCTCGTCCGCGCGGCCCGCGCGCATGCTGTGGCGCGC 1823
QY 121 TCGGAGTGTCCCGCGCGCGCTTGGCTGCTCGCGCGCACAGGACGCGGACGCGCGCT 180
Db 1822 TCGGAGTGTCCCGCGCGCGCTTGGCTGCTCGCGCGCACAGGACGCGGACGCGCGCT 1763
QY 181 CCGCGCGCGCGAGTGTGAGTACTTGGGCGAGATCCACACATCTCCGCTGGCGGCTC 240
Db 1762 CCGCGCGCGCGAGTGTGAGTACTTGGGCGAGATCCACACATCTCCGCTGGCGGCTC 1703
QY 241 AGNAGGACGACCGCGCGCGCACCGGCGACCGCTGTGGTATTCGCGCATGCGGCGCGCTAC 300
Db 1702 AGNAGGACGACCGCGCGCGCACCGGCGACCGCTGTGGTATTCGCGCATGCGGCGCGCTAC 1643
QY 301 AGCTGAGAGATGAATTCCTCTGCTGACAAACCAACGCTGTCTCTGGAAGGGTGTTTG 360
Db 1642 AGCTGAGAGATGAATTCCTCTGCTGACAAACCAACGCTGTCTCTGGAAGGGTGTTTG 1583
QY 361 GAGGAGTGTGTGTGTGTATCAAGGAGTCCACAAATGCTTAAAGAGCTGTCTTCCAGGGA 420
Db 1582 GAGGAGTGTGTGTGTGTATCAAGGAGTCCACAAATGCTTAAAGAGCTGTCTTCCAGGGA 1523
QY 421 GTGAAAATCTGGGATGCGAATGGATCCGAGACTTTTGGACAGCTGGGATTTCTCCACC 480
Db 1522 GTGAAAATCTGGGATGCGAATGGATCCGAGACTTTTGGACAGCTGGGATTTCTCCACC 1463
QY 481 AGAGAGAGGAGGACTTGGGCGCGAGTTTATGGCTTCCAGTGGAGCAATTTGGGGCAGAA 540
Db 1462 AGAGAGAGGAGGACTTGGGCGCGAGTTTATGGCTTCCAGTGGAGCAATTTGGGGCAGAA 1403
QY 541 TACAGAGATATGAATCAGATTATTCAGGACAGGAGTTCACCACTGCAAAAGAGTGATT 600
Db 1402 TACAGAGATATGAATCAGATTATTCAGGACAGGAGTTCACCACTGCAAAAGAGTGATT 1343
QY 601 GACACCATCAAAACCAACCCCTGACGACAGAGAAATCATCATGTGGCTTGGATTCGAAGA 660
Db 1342 GACACCATCAAAACCAACCCCTGACGACAGAGAAATCATCATGTGGCTTGGATTCGAAGA 1283
QY 661 GATCTTCTCTGATGGGCTGCTTCCATGCGCATGCCCTCTGCCAGTTCATGTGGTGAAC 720
Db 1282 GATCTTCTCTGATGGGCTGCTTCCATGCGCATGCCCTCTGCCAGTTCATGTGGTGAAC 1223
QY 721 AGTGAGTGTCTGCTGCGAGCTGTACAGAGATCGGAGACATGGGCTCGGTGCTTTTC 780
Db 1222 AGTGAGTGTCTGCTGCGAGCTGTACAGAGATCGGAGACATGGGCTCGGTGCTTTTC 1163
QY 781 AACATCGCGAGCTACGCCCTCTGCTGACATCATGATTGGCAGATCAACCGGCTGAAGCCA 840
Db 1162 AACATCGCGAGCTACGCCCTCTGCTGACATCATGATTGGCAGATCAACCGGCTGAAGCCA 1103
QY 841 GGTGACTTTATACACATTTTGGGAGATGACATATTTACTTGAATCATCATCGAGCACTG 900
Db 1102 GGTGACTTTATACACATTTTGGGAGATGACATATTTACTTGAATCATCATCGAGCACTG 1043
QY 901 AAAATTCAGCTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGGATTTCTGAAAAGTT 960
Db 1042 AAAATTCAGCTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGGATTTCTGAAAAGTT 983
QY 961 GAGAAAATGATGACTTCAAGCTGAAGACTTTCAGATTTCAAGGGTACAAATCCGCATCCA 1020
Db 982 GAGAAAATGATGACTTCAAGCTGAAGACTTTCAGATTTCAAGGGTACAAATCCGCATCCA 923
QY 1021 ACTATTAATGAAATGGCTTTTGGGCTGCTTTTAAAGAGCTGGAAGATTTGTCA 1080
Db 922 ACTATTAATGAAATGGCTTTTGGGCTGCTTTTAAAGAGCTGGAAGATTTGTCA 863
QY 1081 GTCTTTAGGGGTTGGGCTGGAATGCCAGGATTAAGATTTCTTTTGTCTTAAAGANAAG 1140
```

Db	1382	GATCTTCCTTGATGGGGCTGCTCCATGCGCATGCCCTCTGCCAGTTCTATGTGGTAGAC	1323
Qy	721	AGTGAGCTGTCTCTGCCAGCTGTACAGAGATCGGGAGACATGGGCTCGGTGTCCTTTC	780
Db	1322	AGTGAGCTGTCTCTGCCAGCTGTACAGAGATCGGGAGACATGGGCTCGGTGTCCTTTC	1263
Qy	781	AACATGCGCCAGCTACGCCCTGTCTACGTACATGATTCGGACATCACGGGCTTAAGCCA	840
Db	1262	AACATGCGCCAGCTACGCCCTGTCTACGTACATGATTCGGACATCACGGGCTTAAGCCA	1203
Qy	841	GGTGACTTTATACACACTTTGGGAGATGCACATATTTACCTGAATCACATCGAGCCACTG	900
Db	1202	GGTGACTTTATACACACTTTGGGAGATGCACATATTTACCTGAATCACATCGAGCCACTG	1143
Qy	901	AAAAATTCAGCTTCAGCGAGAACCCAGACCTTCCCAAGCTCAGGATTTCTTCGAAAAGTT	960
Db	1142	AAAAATTCAGCTTCAGCGAGAACCCAGACCTTCCCAAGCTCAGGATTTCTTCGAAAAGTT	1083
Qy	961	GAGAAAAATGTATGATCTCAAAGCTGAAGACTTTCAGATTCAGGGGTACAATCCGCAATCCA	1020
Db	1082	GAGAAAAATGTATGATCTCAAAGCTGAAGACTTTCAGATTCAGGGGTACAATCCGCAATCCA	1023
Qy	1021	ACTATTTAAATGGAATGGCTGTATTAGGGTCTTTCAAAGGCTGAAAGGATATGTGCA	1080
Db	1022	ACTATTTAAATGGAATGGCTGTATTAGGGTCTTTCAAAGGCTGAAAGGATATGTGCA	963
Qy	1081	GTCTTTTAGGGGTTCGGCTGGATCCCGAGGTAAAAAGTTCTTTTTTCCTCTAAAAAGAAAGG	1140
Db	962	GTCTTTTAGGGGTTCGGCTGGATCCCGAGGTAAAAAGTTCTTTTTTCCTCTAGCTGATTCCTG	903
Qy	1141	AACTAGGTCAAAAAATCTGTCCG	1162
Db	902	AGGCTGTATGACTTCTTAGCTG	881

```

RESULT 6
US-09-724-676A-34515/c
; Sequence 34515, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 37222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34515
; LENGTH: 2042
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-34515

```

Query Match	73.3%;	Score 1126.4;	DB 5;	Length 2042;
Best Local Similarity	98.0%;	Prod. No. 2.6e-275;		
Matches 1139;	Conservative	0;	Mismatches 23;	Indels 0; Gaps 0;
QY	1	GGGGGGGGGGACACATTGGCCTGCCTCCGTCGGCGCGGCACATTGGCCTGCCTCCGT	60	
Db	2042	GGGGGGGGGGGACACATTGGCCTGCCTCCGTCGGCGCGGCACATTGGCCTGCCTCCGT	1983	
QY	61	CCGCGCGCGCCACTTCGCTCCGTCGCCCGCGCGGCCACATCGCTGTGGCGCGC	120	
Db	1982	CCGCGCGCGCCACTTCGCTCCGTCGCCCGCGCGGCCACATCGCTGTGGCGCGC	1923	
QY	121	TCGGAGCTTCGCGCGCGCGCCTTGCCCCCGCCGACAGAGCGGGAGCGCGACGCGT	180	
Db	1922	TCGGAGCTTCGCGCGCGCGCCTTGCCCCCGCCGACAGAGCGGGAGCGCGACGCGT	1863	
QY	181	CCGCGCACGGGGAGCTGCAGTACTTGGGGGACGATCCAACACATCTCCGTGCGGCTC	240	
Db	1862	CCGCGCACGGGGAGCTGCAGTACTTGGGGGACGATCCAACACATCTCCGTGCGGCTC	1803	

Db 862 GTCCTTAGGGGTTGGCTGGATGCGAGGTAAAAAGTTCTTTTGTGTTTACGTGAATTCCTG 803

Qy 1141 AACTAGGTCAAAAATCTGTCCG 1162

Db 802 AGGCTGGTATGACTTCTAGCTG 781

RESULT 5

US-09-724-676-34515/c

; Sequence 34515, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: CompuGen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 CompuGen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 34515

; LENGTH: 2042

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-724-676-34515

Query Match 73.3%; Score 1126.4; DB 5; Length 2042;

Best Local Similarity 98.0%; Pred. No. 2.6e-275;

Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCGCGGCGCCACTTGGCTGCTCCCT 60

Db 2042 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCGCGGCGCCACTTGGCTGCTCCCT 1983

Qy 61 CCGCGCGGCCACTTGGCTGCTCCGTCGCGCGCGGCGGCGCCACTTGGCTGCTCCCT 120

Db 1982 CCGCGCGGCCACTTGGCTGCTCCGTCGCGCGCGGCGGCGCCACTTGGCTGCTCCCT 1923

Qy 121 TCGGAGCTCGCGCGCGGCGCTTGGCTGCGCGCGGCGGCGGCGGCGGCGGCGGCT 180

Db 1922 TCGGAGCTCGCGCGCGGCGCTTGGCTGCGCGCGGCGGCGGCGGCGGCGGCGGCT 1863

Qy 181 CCGCGCGCGGGGAGCTGCACTTGGGCGAGATCCAAACATCTCTCGCTGGCGGCTC 240

Db 1862 CCGCGCGCGGGGAGCTGCACTTGGGCGAGATCCAAACATCTCTCGCTGGCGGCTC 1803

Qy 241 AGGAAGGACGACCGCGCGGCGACCGCTGTGCTGATTCGGAATCCAAACATCTCTCGCT 300

Db 1802 AGGAAGGACGACCGCGCGGCGACCGCTGTGCTGATTCGGAATCCAAACATCTCTCGCT 1743

Qy 301 AGCTGAGAGATGAATTCCTCTGCTGACCAACAACTGTGTTCTGGAAGGTTGTTT 360

Db 1742 AGCTGAGAGTGAATTCCTCTGCTGACCAACAACTGTGTTCTGGAAGGTTGTTT 1683

Qy 361 GAGAGTGTGTTGTTTATCAAGGATCCAAATGCTAAAGAGTGTCTTCCAAAGGA 420

Db 1682 GAGAGTGTGTTGTTTATCAAGGATCCAAATGCTAAAGAGTGTCTTCCAAAGGA 1623

Qy 421 GTGAAATCTGGGATGCCAATGGATCCGAGACTTTTGGACAGCTTGGGATTCACAC 480

Db 1622 GTGAAATCTGGGATGCCAATGGATCCGAGACTTTTGGACAGCTTGGGATTCACAC 1563

Qy 481 AGAGAAGAGGGGACTTGGGCGCCAGTTATGGCTCCAGTGGAGGCAATTTGGGCGAGAA 540

Db 1562 AGAGAAGAGGGGACTTGGGCGCCAGTTATGGCTCCAGTGGAGGCAATTTGGGCGAGAA 1503

Qy 541 TACAGATATGGAATTCAGATTAATTCAGGACAGGAGTGTGCAACTGTCAAGAGTGATT 600

Db 1502 TACAGATATGGAATTCAGATTAATTCAGGACAGGAGTGTGCAACTGTCAAGAGTGATT 1443

Qy 601 GACACCATCAAAACCAACCTGACGACAGGAATCATCTGTGCTTGGATCCAAGA 660

Db 1442 GACACCATCAAAACCAACCTGACGACAGGAATCATCATGTGCTTGGATCCAAGA 1383

Qy 661 GATCTTCTGTGATGGCGCTGCTCCATGCCATGCCCTCTGCCAGTTCATATGTGTGTAAC 720

CURRENT FILING DATE: 2000-11-28									
NUMBER OF SEQ ID NOS: 97222									
SOFTWARE: Patentin version 3.2									
SEQ ID NO 34597									
LENGTH: 2391									
TYPE: DNA									
ORGANISM: Homo sapiens									
US-09-724-676-34597									
Query Match 73.3%; Score 1126.4; DB 5; Length 2391;									
Best Local Similarity 98.0%; Pred. No. 2.8e-275;									
Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps									
QY	1	GGGGGGGGGGGGGACCACTTGGCTTCGGTCCCGCGCGCACTTGGCTTCGGT	60						
DB	2391	GGGGGGGGGGGGGACCACTTGGCTTCGGTCCCGCGCGCACTTGGCTTCGGT	2332						
QY	61	CCCGCCGGCCACTTCCCTGCTCCCGTCCCGCGCGCACTTGGCTTCGGT	120						
DB	2331	CCCGCCGGCCACTTCCCTGCTCCCGTCCCGCGCGCACTTGGCTTCGGT	2272						
QY	121	TGGGAGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	180						
DB	2271	TGGGAGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	2212						
QY	181	CCCG	240						
DB	2211	CCCG	2152						
QY	241	AGGAG	300						
DB	2151	AGGAG	2092						
QY	301	AGCTGAGAGATGAATTCCTTCCTGCAACAACGCTGCTTGGAGGGGTGTTG	360						
DB	2091	AGCTGAGAGATGAATTCCTTCCTGCAACAACGCTGCTTGGAGGGGTGTTG	2032						
QY	361	GAGAGATGCTGTGTATTATCAAGGATCCAAATGTAAAGGCTGTCTTCAAGGA	420						
DB	2031	GAGAGATGCTGTGTATTATCAAGGATCCAAATGTAAAGGCTGTCTTCAAGGA	1972						
QY	421	GTTAAATGTGGAGTGGCAATGGATCCGAGACTTTTGGACAGCTGGGATTCACAC	480						
DB	1971	GTTAAATGTGGAGTGGCAATGGATCCGAGACTTTTGGACAGCTGGGATTCACAC	1912						
QY	481	AG	540						
DB	1911	AG	1852						
QY	541	TACAGAGATATGATCAGATTATTACAGAGAGAGAGATGACCACTGCAAGAGTATT	600						
DB	1851	TACAGAGATATGATCAGATTATTACAGAGAGAGAGATGACCACTGCAAGAGTATT	1792						
QY	601	GACACCATCAAAACCAACCTTGACGACAGAGAAATCATATGTGCGTTGAAATCCAGA	660						
DB	1791	GACACCATCAAAACCAACCTTGACGACAGAGAAATCATATGTGCGTTGAAATCCAGA	1732						
QY	661	GATCTTCTCTGATGGCGCTGCTTCATGTCATGCCCTCTGCGCAGTTCTATGTGGTGAAC	720						
DB	1731	GATCTTCTCTGATGGCGCTGCTTCATGTCATGCCCTCTGCGCAGTTCTATGTGGTGAAC	1672						
QY	721	AGTGAAGCTGCTGCGCACTGTCACAGATGGAGAGAGAGAGAGAGAGAGAGAG	780						
DB	1671	AGTGAAGCTGCTGCGCACTGTCACAGATGGAGAGAGAGAGAGAGAGAGAGAG	1612						
QY	781	AACATGCGCAGCTTACCGCTCTCAAGTATGATGGAGAGAGAGAGAGAGAGAGAG	840						
DB	1611	AACATGCGCAGCTTACCGCTCTCAAGTATGATGGAGAGAGAGAGAGAGAGAGAG	1552						
QY	841	GGTGAATTTATACACTTTTGGAGATGACATATTTACCTGAATCAGATGAGCACTG	900						
DB	1551	GGTGAATTTATACACTTTTGGAGATGACATATTTACCTGAATCAGATGAGCACTG	1492						

Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGCACTTGGCTGCTCCGTCGGCGGCACTTGGCTGCTCCGT 60

DB 2411 GGGGGGGGGGGGCACTTGGCTGCTCCGTCGGCGGCACTTGGCTGCTCCGT 2352

QY 61 CCGCGCGCGCACTTGGCTGCTCCGTCGGCGGCACTTGGCTGCTCCGT 120

DB 2351 CCGCGCGCGCACTTGGCTGCTCCGTCGGCGGCACTTGGCTGCTCCGT 2292

QY 121 TCGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180

DB 2291 TCGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2232

QY 181 CCGCGCGCGGAGCTGAGTCACTTGGCGGCACTTGGCGGCACTTGGCGGCGTC 240

DB 2231 CCGCGCGCGGAGCTGAGTCACTTGGCGGCACTTGGCGGCACTTGGCGGCGTC 2172

QY 241 AGGAGGAGCAGCCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCTAC 300

DB 2171 AGGAGGAGCAGCCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCTAC 2112

QY 301 AGCTGAGAGTGAATTCCTCTGCTGCAACAAGGCGTGTCTGGAAGGCTTTTG 360

DB 2111 AGCTGAGAGTGAATTCCTCTGCTGCAACAAGGCGTGTCTGGAAGGCTTTTG 2052

QY 361 GAGGAGTGTGCTGCTGCTTCAAGGAGTCCAAAGTCTTAAAGCTGCTTCAAGGGA 420

DB 2051 GAGGAGTGTGCTGCTGCTTCAAGGAGTCCAAAGTCTTAAAGCTGCTTCAAGGGA 1992

QY 421 GTGAAATCTGGGATGCCAATGGATCCGAGACTTTTGGACAGGCTGGGATTTCTCAC 480

DB 1991 GTGAAATCTGGGATGCCAATGGATCCGAGACTTTTGGACAGGCTGGGATTTCTCAC 1932

QY 481 AGGAGAGAGGGGAGCTTGGGAGCTTGGGATTTGGCTTGGGAGGAGCTTTGGGAGAA 540

DB 1931 AGGAGAGAGGGGAGCTTGGGAGCTTGGGATTTGGGAGCTTGGGAGGAGCTTTGGGAGAA 1872

QY 541 TACAGAGATATGGATCAGATTATTCAGGACAGGAGGATTCAGCACTGCAAGAGTAT 600

DB 1871 TACAGAGATATGGATCAGATTATTCAGGACAGGAGGATTCAGCACTGCAAGAGTAT 1812

QY 601 GACACATCAAAAACCAACCTTGAAGCAAGAGATCATCATGTGGCTTGAATTCAGGA 660

DB 1811 GACACATCAAAAACCAACCTTGAAGCAAGAGATCATCATGTGGCTTGAATTCAGGA 1752

QY 661 GATTTCTCTGATGAGCGCTGCTTCAAGGAGCTTGGCAAGGCTTGAATTCAGGA 720

DB 1751 GATTTCTCTGATGAGCGCTGCTTCAAGGAGCTTGGCAAGGCTTGAATTCAGGA 1692

QY 721 AGTGAAGCTGCTGCGAGCTGACAGAGATCCGAGAGCAATGGGCTCGGATTCCTTTC 780

DB 1691 AGTGAAGCTGCTGCGAGCTGACAGAGATCCGAGAGCAATGGGCTCGGATTCCTTTC 1632

QY 781 AACATCGCAGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840

DB 1631 AACATCGCAGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1572

QY 841 GGTGACTTTATACACTTTGGGAGATGCAATATTTACCTGATTCAGATGAGCGCACTG 900

DB 1571 GGTGACTTTATACACTTTGGGAGATGCAATATTTACCTGATTCAGATGAGCGCACTG 1512

QY 901 AAAATTCAGCTTACGAGGAGACCAAGACTTTTCCAAAGCTGAGATTTCTTCAAAAGTT 960

DB 1511 AAAATTCAGCTTACGAGGAGACCAAGACTTTTCCAAAGCTGAGATTTCTTCAAAAGTT 1452

QY 961 GAGGAAATGATGACTTCAAGGAGAGCACTTCAAGGAGAGCACTTCAAGGAGAGCACT 1020

DB 1451 GAGGAAATGATGACTTCAAGGAGAGCACTTCAAGGAGAGCACTTCAAGGAGAGCACT 1392

QY 1021 ACTATTAAATGGAATGGCTGCTTGAAGGCTTCAAGGAGAGCACTTGAAGGAGTATTCGA 1080

DB 1391 ACTATTAAATGGAATGGCTGCTTGAAGGCTTCAAGGAGAGCACTTGAAGGAGTATTCGA 1332

QY 1081 GTCTTTAGGGGTTGGGCTGATGCGGAGTAAAGTCTTTTGTCTTAAAGAAAGG 1140

DB 1331 GTCTTTAGGGGTTGGGCTGATGCGGAGTAAAGTCTTTTGTCTTAAAGTATTCG 1272

QY 1141 AACTAGGTCAAAATCTGTCG 1162

DB 1271 AGGGGTATGACTTCTAGCTG 1250

RESULT 10

US-09-724-676A-34535/c

; Sequence 34535, Application US/09724676A

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676A

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 34535

; LENGTH: 2411

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (41)-(41)

; OTHER INFORMATION: n is a,c,g, or t

US-09-724-676A-34535

Query Match 73.3%; Score 1126.4; DB 5; Length 2411;

Best Local Similarity 98.0%; Pred. No. 2,8e-275;

Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGCACTTGGCTGCTCCGTCGGCGGCACTTGGCTGCTCCGT 60

DB 2411 GGGGGGGGGGGGCACTTGGCTGCTCCGTCGGCGGCACTTGGCTGCTCCGT 2352

QY 61 CCGCGCGCGCACTTGGCTGCTCCGTCGGCGGCACTTGGCTGCTCCGT 120

DB 2351 CCGCGCGCGCACTTGGCTGCTCCGTCGGCGGCACTTGGCTGCTCCGT 2292

QY 121 TCGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180

DB 2291 TCGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2232

QY 181 CCGCGCGCGGAGCTGAGTCACTTGGCGGCACTTGGCGGCACTTGGCGGCGTC 240

DB 2231 CCGCGCGCGGAGCTGAGTCACTTGGCGGCACTTGGCGGCACTTGGCGGCGTC 2172

QY 241 AGGAGGAGCAGCCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCTAC 300

DB 2171 AGGAGGAGCAGCCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCTAC 2112

QY 301 AGCTGAGAGTGAATTCCTCTGCTGCAACAAGGCGTGTCTGGAAGGCTTTTG 360

DB 2111 AGCTGAGAGTGAATTCCTCTGCTGCAACAAGGCGTGTCTGGAAGGCTTTTG 2052

QY 361 GAGGAGTGTGCTGCTGCTTCAAGGAGTCCAAAGTCTTAAAGCTGCTTCAAGGGA 420

DB 2051 GAGGAGTGTGCTGCTGCTTCAAGGAGTCCAAAGTCTTAAAGCTGCTTCAAGGGA 1992

QY 421 GTGAAATCTGGGATGCCAATGGATCCGAGACTTTTGGACAGGCTTGAATTCAGGA 480

DB 1991 GTGAAATCTGGGATGCCAATGGATCCGAGACTTTTGGACAGGCTTGAATTCAGGA 1932

QY 481 AGGAGAGAGGGGAGCTTGGGAGCTTGGGATTTGGGAGGAGCTTTGGGAGAA 540

DB 1931 AGGAGAGAGGGGAGCTTGGGAGCTTGGGATTTGGGAGGAGCTTTGGGAGAA 1872

QY 541 TACAGAGATATGGATCAGATTATTCAGGACAGGAGTTCAGCACTGCAAGAGTAT 600

US-09-724-676A-34589/c
 ; Sequence 34589, Application US/09724676A
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676A
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 34589
 ; LENGTH: 2491
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-724-676A-34589

Query Match 73.3%; Score 1126.4; DB 5; Length 2491;
 Best Local Similarity 98.0%; Pred. No. 2.9e-275;
 Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGACCACTTGCTGCTCCGCGCGCGCACTTGAGCTGCTCCGCT 60
 Db 2491 GGGGGGGGGGGGACCACTTGCTGCTCCGCTCCGCGCGCGCACTTGAGCTGCTCCGCT 2432
 QY 61 CCGCGCGCGCACTTGCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 Db 2431 CCGCGCGCGCACTTGCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCG 2372
 QY 121 TCGAGCTCCCG 180
 Db 2371 TCGAGCTCCCG 2312
 QY 181 CCG 240
 Db 2311 CCG 2252
 QY 241 AGGAAGAGGAGCG 300
 Db 2251 AGGAAGAGGAGCG 2192
 QY 301 AGCTTGAAGATGATTCCTGCTGCTGCAACCAACGCTGCTTGGAAAGGCTGTTG 360
 Db 2191 AGCTTGAAGATGATTCCTGCTGCTGCTGCAACCAACGCTGCTTGGAAAGGCTGTTG 2132
 QY 361 GAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 Db 2131 GAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2072
 QY 421 GTGAAATCTGGAGATGCAATGATCCGAGACTTTTGGAGAGCTGGGATTCCTGACC 480
 Db 2071 GTGAAATCTGGAGATGCAATGATCCGAGACTTTTGGAGAGCTGGGATTCCTGACC 2012
 QY 481 AGGAAGAGGAGGAGCTGGGCGCGCTTGAAGCTTCCAGTGAAGGAGCAATTTGGGGAGAA 540
 Db 2011 AGGAAGAGGAGGAGCTGGGCGCGCTTGAAGCTTCCAGTGAAGGAGCAATTTGGGGAGAA 1952
 QY 541 TACAGAGATATGAGATGATTTATTCAGGACAGGAGGTTGACCAACTGCAAAAGTGATT 600
 Db 1951 TACAGAGATATGAGATGATTTATTCAGGACAGGAGGTTGACCAACTGCAAAAGTGATT 1892
 QY 601 GACACCATCAAAACCAACCTTGAAGAGAGAGATATGATGAGCTTGGAAATCCAAAG 660
 Db 1891 GACACCATCAAAACCAACCTTGAAGAGAGAGATATGATGAGCTTGGAAATCCAAAG 1832
 QY 661 GATCTTCTCTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 Db 1831 GATCTTCTCTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1772
 QY 721 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 Db 1771 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1712
 QY 781 AACATGCGAGCTAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840

Db 1711 AACATGCGAGCTAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1652
 QY 841 GGTGACTTATACACACTTGGGAGATGACATATTTACTGATACATGAGCGCACTG 900
 Db 1651 GGTGACTTATACACACTTGGGAGATGACATATTTACTGATACATGAGCGCACTG 1592
 QY 901 AAAATTCAGCTTACAGCGAGAACCCAGACTTTCACAAAGCTCAGGATTTCTGAAAAGTT 960
 Db 1591 AAAATTCAGCTTACAGCGAGAACCCAGACTTTCACAAAGCTCAGGATTTCTGAAAAGTT 1532
 QY 961 GAGAAATATGATGATTCGAAAGCTGAAAGCTTTCAGATTGAAGGATCAATCCGATCCA 1020
 Db 1531 GAGAAATATGATGATTCGAAAGCTGAAAGCTTTCAGATTGAAGGATCAATCCGATCCA 1472
 QY 1021 ACTATTAATATGAAATGCTGTTTGAAGGCTTTCAGAAAGGCTTTCAGAAAGGCTTTC 1080
 Db 1471 ACTATTAATATGAAATGCTGTTTGAAGGCTTTCAGAAAGGCTTTCAGAAAGGCTTTC 1412
 QY 1081 GTCTTTAAGGCTTGGGCTGAGATGCGAGGTAAGTTCCTTTTGTCTTAAAGAAAGG 1140
 Db 1411 GTCTTTAAGGCTTGGGCTGAGATGCGAGGTAAGTTCCTTTTGTCTTAAAGGATTCCTG 1352
 QY 1141 AACTAGTCAAAAATCTGTCCG 1162
 Db 1351 AAGGTGATATGACTTCTAGCTG 1330

RESULT 13

US-09-724-676-34548/c
 ; Sequence 34548, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 34548
 ; LENGTH: 2527
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (41)..(41)
 ; OTHER INFORMATION: n is a,c,g, or t
 US-09-724-676-34548

Query Match 73.3%; Score 1126.4; DB 5; Length 2527;
 Best Local Similarity 98.0%; Pred. No. 2.9e-275;
 Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGACCACTTGCTGCTCCGCGCGCGCACTTGAGCTGCTCCGCT 60
 Db 2527 GGGGGGGGGGGGACCACTTGCTGCTCCGCGCGCGCGCACTTGAGCTGCTCCGCT 2468
 QY 61 CCGCGCGCGCACTTGCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 Db 2467 CCGCGCGCGCACTTGCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCG 2408
 QY 121 TCGAGCTGCG 180
 Db 2407 TCGAGCTGCG 2348
 QY 181 CCGCGCGCGAGCTGCACTTCTGAGGAGAGATCAACATCTCTCGCTGCGGCTC 240
 Db 2347 CCGCGCGCGAGCTGCACTTCTGAGGAGAGATCAACATCTCTCGCTGCGGCTC 2288
 QY 241 AGGAAGAGGAGCG 300
 Db 2287 AGGAAGAGGAGCG 2228


```

QY 901 AAAATTAGCTTACGAGAACCCGACCTTCCCAAGTCGAGATCTTCCGAAGTT 960
DB 1627 AAAATTACGCTTACGAGAACCCGACCTTCCCAAGTCGAGATCTTCCGAAGTT 1568
QY 961 GAGAAATTTATGATCTTCAAGCTGAAGATCTTCAATGAGAGGATCAATCCCATCCA 1020
DB 1567 GAGAAATTTATGATCTTCAAGCTGAAGATCTTCAATGAGAGGATCAATCCCATCCA 1508
QY 1021 ACTATTAATGAAATGAGCTGTTAGAGTGCTTTCAAGAGCTTGAAGATTAATGCA 1080
DB 1507 ACTATTAATGAAATGAGCTGTTAGAGTGCTTTCAAGAGCTTGAAGATTAATGCA 1448
QY 1081 GTCTTTAGGGGTTGGCTGGATGCCAGATAAAGTTCTTTTGTCTTAAAGAAAG 1140
DB 1447 GTCTTTAGGGGTTGGCTGGATGCCAGATAAAGTTCTTTTGTCTTAAAGAAAG 1388
QY 1141 AACTAGGTCAAAAATCTGTCCG 1162
DB 1387 AGGGGTGATGACTTCTAGCTG 1366

RESULT 15
US-09-724-676-34565/c
; Sequence 34565, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34565
; LENGTH: 3023
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-34565

Query Match          73.3%; Score 1126.4; DB 5; Length 3023;
Best Local Similarity 98.0%; Pred. No. 3.1e-275;
Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGACCACTTGGCTTGCCTCCCGCGCGGCACTTGGCTTGCCTCCGT 60
DB 3023 GGGGGGGGGGGGACCACTTGGCTTGCCTCCCGCGCGGCACTTGGCTTGCCTCCGT 2964
QY 61 CCCGCGCGGCACTTGGCTTGCCTCCCGCGCGGCACTTGGCTTGCCTCCGT 120
DB 2963 CCCGCGCGGCACTTGGCTTGCCTCCCGCGCGGCACTTGGCTTGCCTCCGT 2904
QY 121 TCGAGAGTGCAGCGCGGCGCTTGCCTCCCGCGCGGCACTTGGCTTGCCTCCGT 180
DB 2903 TCGAGAGTGCAGCGCGGCGCTTGCCTCCCGCGCGGCACTTGGCTTGCCTCCGT 2844
QY 181 CCGCGCGAGCGGAGCTGAGTACTTGGGCGAGATCCACATCTCTCCGTCGCGCTC 240
DB 2843 CCGCGCGAGCGGAGCTGAGTACTTGGGCGAGATCCACATCTCTCCGTCGCGCTC 2784
QY 241 AGAAGAGCAGCCGACCGGCGACCGGCACTTGTGAGTATTCGAGATGCAAGCGCTAC 300
DB 2783 AGAAGAGCAGCCGACCGGCGACCGGCACTTGTGAGTATTCGAGATGCAAGCGCTAC 2724
QY 301 AGCTTGAGAGTAAATTCCTCTGCTGACAAACAACTGTGTCTTGAAGGGTGTTTTG 360
DB 2723 AGCTTGAGAGTAAATTCCTCTGCTGACAAACAACTGTGTCTTGAAGGGTGTTTTG 2664
QY 361 GAGGAGTGTGCTGTGTTTTCAGAGGATCCCAATGTCTAAAGAGCTGTCTTCCAGGGA 420
DB 2663 GAGGAGTGTGCTGTGTTTTCAGAGGATCCCAATGTCTAAAGAGCTGTCTTCCAGGGA 2604
QY 421 GTGAAATCTGAGATGCCAATGATCCGAGACTTTTGGACAGCTGAGATTTCCACG 480
DB 2603 GTGAAATCTGAGATGCCAATGATCCGAGACTTTTGGACAGCTGAGATTTCCACG 2544

```

```

QY 481 AGAAGAAAGGGGACTTGGGCCCAAGTTATGCTTCAATGAGGCAATTTTGGGCGAGA 540
DB 2543 AGAAGAAAGGGGACTTGGGCCCAAGTTATGCTTCAATGAGGCAATTTTGGGCGAGA 2484
QY 541 TACAGAGATATGGAATCAAGTATTAATGAGGAGAGGATGACCAATGCAAGATGAT 600
DB 2483 TACAGAGATATGGAATCAAGTATTAATGAGGAGAGGATGACCAATGCAAGATGAT 2424
QY 601 GACACATCAAAACCAACCTGACGACGAGAGATATATATGTGGCTTGAATCCAAAG 660
DB 2423 GACACATCAAAACCAACCTGACGACGAGAGATATATATGTGGCTTGAATCCAAAG 2364
QY 661 GATTTCTCTGATGAGGCTGCTCCATGCCATGCCCTGCACTTGTATGATGATGAG 720
DB 2363 GATTTCTCTGATGAGGCTGCTCCATGCCATGCCCTGCACTTGTATGATGATGAG 2304
QY 721 AGTGAAGTGTCTTGCAGCTGATCCAGAGATCCGAGAGATGAGGCTCGATGCTTTG 780
DB 2303 AGTGAAGTGTCTTGCAGCTGATCCAGAGATCCGAGAGATGAGGCTCGATGCTTTG 2244
QY 781 AACATGCGCAGTAAAGCCTGCTCAGTACATGATTCGACATCAAGGCTTGAAGCCA 840
DB 2243 AACATGCGCAGTAAAGCCTGCTCAGTACATGATTCGACATCAAGGCTTGAAGCCA 2184
QY 841 GGTGACTTTATACACACTTGGAGAGTGCACATATTTACTGATACATGAGCCACTG 900
DB 2183 GGTGACTTTATACACACTTGGAGAGTGCACATATTTACTGATACATGAGCCACTG 2124
QY 901 AAAATTCAGCTTACGAGAACCCGACCTTCCCAAGCTCAGAGATTTCTGAAAAGTT 960
DB 2123 AAAATTCAGCTTACGAGAACCCGACCTTCCCAAGCTCAGAGATTTCTGAAAAGTT 2064
QY 961 GAGAAATTTATGATCTTCAAAAGCTGAAGATTTGAAGGGTGAATCCGATCCA 1020
DB 2063 GAGAAATTTATGATCTTCAAAAGCTGAAGATTTGAAGGGTGAATCCGATCCA 2004
QY 1021 ACTATTAATGAAATGAGCTGTTAGAGTGCTTTCAAGAGCTTGAAGATTAATGCA 1080
DB 2003 ACTATTAATGAAATGAGCTGTTAGAGTGCTTTCAAGAGCTTGAAGATTAATGCA 1944
QY 1081 GTCTTTAGGGGTTGGCTGGATGCCAGATAAAGTTCTTTTGTCTTAAAGAAAG 1140
DB 1943 GTCTTTAGGGGTTGGCTGGATGCCAGATAAAGTTCTTTTGTCTTAAAGAAAG 1884
QY 1141 AACTAGGTCAAAAATCTGTCCG 1162
DB 1883 AACTAGGTCAAAAATCTGTCCG 1862

```

Search completed: November 29, 2002, 00:20:06
Job time : 123.086 secs


```

22 1505.4 98.0 1829 29 US-09-726-806-5390 Sequence 5390, Ap
23 1505.4 98.0 1829 29 US-09-726-807-3895 Sequence 3895, Ap
24 1493.4 97.2 3298 63 US-60-195-106-200 Sequence 200, App
25 1493.4 97.2 3298 63 US-60-217-674-28 Sequence 28, App1
26 1493.4 97.2 3298 71 US-60-278-258-674 Sequence 674, App
27 1493 97.2 1540 1 PCT-US01-14827-3952 Sequence 3952, Ap
28 1493 97.2 1540 22 US-09-577-408-4887 Sequence 4887, Ap
29 1478.4 96.2 1633 61 US-60-172-573-69 Sequence 69, App
30 1422.2 92.6 1478 39 US-10-084-817-332 Sequence 332, App
31 1360 88.5 1721 30 US-09-760-475-391 Sequence 391, App
32 1325 86.3 1395 30 US-09-760-475-1516 Sequence 1516, App
33 1325 86.3 1395 30 US-09-760-485-46 Sequence 46, App1
34 1325 86.3 1395 42 US-10-216-436-46 Sequence 46, App1
35 1280 83.3 1539 1 PCT-US01-08631-20764 Sequence 20764, A
36 942 61.3 942 18 US-09-457-205-29 Sequence 29, App
37 942 61.3 942 24 US-09-631-275-143 Sequence 143, App
38 612 39.8 18597 25 US-09-658-659-8 Sequence 8, App1
39 612 39.8 18597 36 US-09-962-665-8 Sequence 8, App1
40 612 39.8 18597 36 US-09-962-677-8 Sequence 8, App1
41 612 39.8 18597 36 US-09-963-333-8 Sequence 8, App1
42 611.8 39.8 18596 22 US-09-577-266-11 Sequence 11, App1
43 611.8 39.8 18596 24 US-09-631-275-140 Sequence 140, App
44 611.8 39.8 18596 33 US-09-880-107-1590 Sequence 1590, App
45 611.8 39.8 18596 36 US-09-954-531-124 Sequence 124, App

```

ALIGNMENTS

RESULT 1

```

PCT-US02-18947-556
: Sequence 556, Application PC/TUS0218947
: GENERAL INFORMATION:
: APPLICANT: Rosetta Inpharmatics
: TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
: FILE REFERENCE: 9301-175-228
: CURRENT APPLICATION NUMBER: PCT/US02/18947
: CURRENT FILING DATE: 2002-06-14
: PRIOR APPLICATION NUMBER: 60/380,770
: PRIOR FILING DATE: 2002-05-14
: NUMBER OF SEQ ID NOS: 2699
: SEQ ID NO 556
: LENGTH: 1536
: TYPE: DNA
: ORGANISM: Homo sapiens
: PUBLICATION INFORMATION:
: DATABASE ACCESSION NUMBER: NM_001071
: DATABASE ENTRY DATE: 2001-06-18
PCT-US02-18947-556

```

```

Query Match          99.8%; Score 1533; DB 1; Length 1536;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 GGGGGGGGGGGGACCTTGGCTGCTCCGTCCGCGCGGCACTTGGCTGCTCCGT 60
Db 1 GGGGGGGGGGGGACCTTGGCTGCTCCGTCCGCGCGGCACTTGGCTGCTCCGT 60
QY 61 CCGCGCGGCACTTGGCTGCTCCGTCCGCGCGGCACTTGGCTGCTCCGT 120
Db 61 CCGCGCGGCACTTGGCTGCTCCGTCCGCGCGGCACTTGGCTGCTCCGT 120
QY 121 TCGGAGCTGCGCGCGGCTTGGCTCCGCGCGGCACTTGGCTGCTCCGT 180
Db 121 TCGGAGCTGCGCGCGGCTTGGCTCCGCGCGGCACTTGGCTGCTCCGT 180
QY 181 CCGCGCGGCACTTGGCTGCTCCGTCCGCGCGGCACTTGGCTGCTCCGT 240
Db 181 CCGCGCGGCACTTGGCTGCTCCGTCCGCGCGGCACTTGGCTGCTCCGT 240
QY 241 AGGAAAGAGGAGCCGAGCGGAGCCGAGCGGAGCCGAGCGGAGCCGAGCGG 300
Db 241 AGGAAAGAGGAGCCGAGCGGAGCCGAGCGGAGCCGAGCGGAGCCGAGCGG 300

```

```

QY 301 AGCGTGAAGATGATTCCTCTGCTGACCAACCAAGTGTCTGGAAGGTGTTTG 360
Db 301 AGCGTGAAGATGATTCCTCTGCTGACCAACCAAGTGTCTGGAAGGTGTTTG 360
QY 361 GAGAGTGTCTGTGTTTATCAAGGATCCAAATGCTAAAGGCTGTCTTCAAGGA 420
Db 361 GAGAGTGTCTGTGTTTATCAAGGATCCAAATGCTAAAGGCTGTCTTCAAGGA 420
QY 421 GTGAAATCTGGATGCGAATGATTCGAGACTTTTGGACAGCTGGGATTCGACC 480
Db 421 GTGAAATCTGGATGCGAATGATTCGAGACTTTTGGACAGCTGGGATTCGACC 480
QY 481 AAGAAAGAGGAGCTTGGGCGCAATTTATGCTTCCAGTGGAGCAATTTGGGAGAA 540
Db 481 AAGAAAGAGGAGCTTGGGCGCAATTTATGCTTCCAGTGGAGCAATTTGGGAGAA 540
QY 541 TACAGATATGATGATGATATTCAGAGCAGGAGTTGACCAATGCAAGAGTGAAT 600
Db 541 TACAGATATGATGATGATATTCAGAGCAGGAGTTGACCAATGCAAGAGTGAAT 600
QY 601 GACACCATCAAAACCAACCTTGAAGCAGAGAAATCATATGTCGCTTGAATCCAGA 660
Db 601 GACACCATCAAAACCAACCTTGAAGCAGAGAAATCATATGTCGCTTGAATCCAGA 660
QY 661 GATCTTCTCTGATGAGGCTGCTCCATGTCATGCTTCCGCTTCAATGTGTGAAC 720
Db 661 GATCTTCTCTGATGAGGCTGCTCCATGTCATGCTTCCGCTTCAATGTGTGAAC 720
QY 721 AGTGACCTGCTCTGCGAGCTGCTTCCAGAGTGGAGAGCATGAGGCTGCTTTC 780
Db 721 AGTGACCTGCTCTGCGAGCTGCTTCCAGAGTGGAGAGCATGAGGCTGCTTTC 780
QY 781 AACATGCCAGCTACGCCCTGCTCAAGTATGATGAGCAGATCAGGAGCTGAGACCA 840
Db 781 AACATGCCAGCTACGCCCTGCTCAAGTATGATGAGCAGATCAGGAGCTGAGACCA 840
QY 841 GGTGATTTATACACTTGTGGAGATGACATATTTACTGAAATCAATGAGCCACTG 900
Db 841 GGTGATTTATACACTTGTGGAGATGACATATTTACTGAAATCAATGAGCCACTG 900
QY 901 AAAATTCAGCTTACGAGGAAACCAAGCTTCCAAAGCTCAGATTTCTGAAAGTT 960
Db 901 AAAATTCAGCTTACGAGGAAACCAAGCTTCCAAAGCTCAGATTTCTGAAAGTT 960
QY 961 GAGAAATGATGATCTTAAAGCTGAAAGCTTCAATGAGGCTGCAATCCGATCCA 1020
Db 961 GAGAAATGATGATCTTAAAGCTGAAAGCTTCAATGAGGCTGCAATCCGATCCA 1020
QY 1021 ACTATTAATGAAATGAGCTGTTAGGCTTCAAGAGCTGAGATATTTGCA 1080
Db 1021 ACTATTAATGAAATGAGCTGTTAGGCTTCAAGAGCTGAGATATTTGCA 1080
QY 1081 GTCTTAAAGGCTTGGCTGATGAGTCCGAGGTAAGTCTTTTGTCTTAAAGAAAG 1140
Db 1081 GTCTTAAAGGCTTGGCTGATGAGTCCGAGGTAAGTCTTTTGTCTTAAAGAAAG 1140
QY 1141 AACTAGGTCAAAATCTGCTGAGACCTATCACTTATTAATTTTAAGATGTCCACT 1200
Db 1141 AACTAGGTCAAAATCTGCTGAGACCTATCACTTATTAATTTTAAGATGTCCACT 1200
QY 1201 GGCATATGTAATCTGCGAGTCTTTTCCAAATTAAGCTTGAATTAATCACTGAGG 1260
Db 1201 GGCATATGTAATCTGCGAGTCTTTTCCAAATTAAGCTTGAATTAATCACTGAGG 1260
QY 1261 GTATCTGAACAATGCTGAGTTATGAACAAGTGAAGATGAATGATGTGCTTGA 1320
Db 1261 GTATCTGAACAATGCTGAGTTATGAACAAGTGAAGATGAATGATGTGCTTGA 1320
QY 1321 CAAAAACATGATGAGTCAATTCATCCAGTACTTATTAAGAGTTGGATTTAC 1380
Db 1321 CAAAAACATGATGAGTCAATTCATCCAGTACTTATTAAGAGTTGGATTTAC 1380

```


QY	1381	AAGCTATTTTGGATATATTTTGAATAATTTTAAAGATTTTCAAGCTATTCCCTCAAA	1440
Db	1381	AAGCTATTTTGGATATATTTTGAATAATTTTAAAGATTTTCAAGCTATTCCCTCAAA	1440
QY	1441	CTGAGGAGCTCGAGTAAACCATTCGATCATGCTTGAAGTGTGTTATGAACCTTTAAAGT	1500
Db	1441	CTGAGGAGCTCGAGTAAACCATTCGATCATGCTTGAAGTGTGTTATGAACCTTTATAGT	1500
QY	1501	TGTTTTATATGTTGCTATAATAAAGAGTGTTCCTGC	1536
Db	1501	TGTTTTATATGTTGCTATAATAAAGAGTGTTCCTGC	1536

RESULT 2

PCT-US99-12080-1
: Sequence 1, Application PC/TUS9912080
: GENERAL INFORMATION:
: APPLICANT: Dean, Nicholas M.
: TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
: TITLE OF INVENTION: HUMAN THYMIDYLATE SYNTHASE EXPRESSION
: NUMBER OF SEQUENCES: 28
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Law Offices of Jane Massey Licata
: STREET: 66 East Main Street
: CITY: Marlton
: STATE: NJ
: COUNTRY: USA
: ZIP: 08053
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: WINDOWS 95
: SOFTWARE: WORD PERFECT 6.1 FOR WINDOWS
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US99/12080
: FILING DATE: herewith
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/089,195
: FILING DATE: June 2, 1998
: ATTORNEY/AGENT INFORMATION:
: NAME: Jane Massey Licata
: REGISTRATION NUMBER: 32,257
: REFERENCE/DOCKET NUMBER: ISPH-0369
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (609) 810-1515
: TELEFAX: (609) 810-1454
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1536 base pairs
: TYPE: Nucleic Acid
: STRANDEDNESS: Single
: TOPOLOGY: Unknown
: PUBLICATION INFORMATION:
: AUTHORS: Takeishi, K.
: AUTHORS: Kaneda, S.
: AUTHORS: Ayusawa, D.
: AUTHORS: Shimizu, K.
: AUTHORS: Gotch, O.
: AUTHORS: Seno, T.
: TITLE: Nucleotide sequence of a functional
: TITLE: cDNA for human thymidylate synthase
: JOURNAL: Nucleic Acids Res.
: VOLUME: 13
: ISSUE: 6
: PAGES: 2035-2043
: DATE: 25-MAR-1985

```
Query Match      99.8%; score 1533; DB 1; Length 1536;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

QY	1	GGGGGGGGGGGACCACTTGGCCTGCTCCGTCGGCGGCGCCACTTGGCGCTGCTCGGT	60
Db	1	GGGGGGGGGGGACCACTTGGCGTGCCTCCGTCGGCGGCGCCACTTGGCGCTGCTCGGT	60
QY	61	CCCGCGCGCCACCTTTCGCGCTGCTCCGTCGGCGGCGCCATGCTGTGGCGCGCG	120
Db	61	CCCGCGCGCGCACTTCGCGCTGCTCCGTCGGCGGCGCCATGCTGTGGCGCGCG	120
QY	121	TCGGAGCTCGCGCGCGCCTTTGCCCGCCCGCCGCGCGCCATGCTGTGGCGCGCG	180
Db	121	TCGGAGCTCGCGCGCGCCTTTGCCCGCCCGCCGCGCCATGCTGTGGCGCGCG	180
QY	181	CCGCGCGACGGGGAGCTGCAGTACCTTGGGGGAGATCCAAACATCTCCGCTCGCGCTC	240
Db	181	CCGCGCGACGGGGAGCTGCAGTACCTTGGGGGAGATCCAAACATCTCCGCTCGCGCTC	240
QY	241	AGGAAGGACGACCGGACCGGGCACCTGTGCGGTATTCGGCATGCGAGGCGCGCTAC	300
Db	241	AGGAAGGACGACCGGACCGGGCACCTGTGCGGTATTCGGCATGCGAGGCGCGCTAC	300
QY	301	AGCCTGAGAGATGAATTCCTCTGCTGACAAACAGAGTGTCTTGGAGAGGTGTTTG	360
Db	301	AGCCTGAGAGATGAATTCCTCTGCTGACAAACAGAGTGTCTTGGAGAGGTGTTTG	360
QY	361	GAGAGTTGCTGTGTTATCAAGGGATCCAAATGCTTAAAGAGCTCTTCCAAAGGGA	420
Db	361	GAGAGTTGCTGTGTTATCAAGGGATCCAAATGCTTAAAGAGCTCTTCCAAAGGGA	420
QY	421	GTGAAATCTGGGATGCCAATGATCCCGAGACTTTTGGACAGGCTGGGATTCCTCCAC	480
Db	421	GTGAAATCTGGGATGCCAATGATCCCGAGACTTTTGGACAGGCTGGGATTCCTCCAC	480
QY	481	AGAGAAGAGGGGACCTTGGGCGCAGTTTATGGCTTCCAGTGGAGGACATTTTGGGGCAGAA	540
Db	481	AGAGAAGAGGGGACCTTGGGCGCAGTTTATGGCTTCCAGTGGAGGACATTTTGGGGCAGAA	540
QY	541	TACAGAGATATCGAATCAGATTTATCAGGACGAGGAGTTGACCAATGCGAAGAGTATT	600
Db	541	TACAGAGATATCGAATCAGATTTATCAGGACGAGGAGTTGACCAATGCGAAGAGTATT	600
QY	601	GACACCATAAACCAACCTGTGACGACAGAGATCATCATGTGCGCTTGGATCCACAG	660
Db	601	GACACCATAAACCAACCTGTGACGACAGAGATCATCATGTGCGCTTGGATCCACAG	660
QY	661	GATCTTCTCTGATGGCGTGCCTCATGSCCATGCCCTCTGCCAGTCTCATGTGTTGAAAC	720
Db	661	GATCTTCTCTGATGGCGTGCCTCATGSCCATGCCCTCTGCCAGTCTCATGTGTTGAAAC	720
QY	721	AGTGAGCTGTCTGCCAGCTGTACAGAGATCGGGAGACATGGGCTCGGTGTGCTTTTC	780
Db	721	AGTGAGCTGTCTGCCAGCTGTACAGAGATCGGGAGACATGGGCTCGGTGTGCTTTTC	780
QY	781	AACATCGCCAGCTAGCGCTGCTCAGCTACATGATTGCGCACATCAOGGGCGTGAAGCCA	840
Db	781	AACATCGCCAGCTAGCGCTGCTCAGCTACATGATTGCGCACATCAOGGGCGTGAAGCCA	840
QY	841	GGTGACTTTATACACACTTTGGGAGATGCCATATTTACCTGAATCAGATCGAGCCACTG	900
Db	841	GGTGACTTTATACACACTTTGGGAGATGCCATATTTACCTGAATCAGATCGAGCCACTG	900
QY	901	AAAAATTCAGCTTACGAGAACCCAGACCTTCCCAAGCTCAGGATCTTCCGAAAGTT	960
Db	901	AAAAATTCAGCTTACGAGAACCCAGACCTTCCCAAGCTCAGGATCTTCCGAAAGTT	960
QY	961	GAGAAATTCATGATCTCAAGCTGAAGACCTTCCAGATGAAAGGTCATATCCGATCCA	1020
Db	961	GAGAAATTCATGATCTCAAGCTGAAGACCTTCCAGATGAAAGGTCATATCCGATCCA	1020
QY	1021	ACTATTAATAATGGAATGGCTTTTAGGGTGCTTTCAAAGGAGCTNGAAGGATTTGTCA	1080
Db	1021	ACTATTAATAATGGAATGGCTTTTAGGGTGCTTTCAAAGGAGCTNGAAGGATTTGTCA	1080
QY	1081	GTCTTTTAGGGGTGGCTGGATGCCGAGGTAAAAAGTTCTTTTGTCTTAAAGAAANAAGG	1140

Db	1081	GTCTTTAGGGTGGGCTGGATGCCGAGGTAAAGTCTCTTTTCCTCTAAAGAAAAGG	1140		241	AGGAGGACGACCGCACGGGCACTGGCATCGCGTATTCGGCATCGAGCGCGCTAC	300	
Qy	1141	AACCTAGTCAAAATCTGTCGGTCACTACGATTAATTAATTTTAAAGGATGTGCCACT	1200		301	AGCTGAGAGATGAATTCCTCTGCTGACACAAACGCTGTGCTGGAAGGCTGTTTG	360	
Db	1141	AACCTAGTCAAAATCTGTCGGTCACTACGATTAATTAATTTTAAAGGATGTGCCACT	1200		301	AGCTGAGAGATGAATTCCTCTGCTGACACAAACGCTGTGCTGGAAGGCTGTTTG	360	
Qy	1201	GGCAAAATGTAATGTCGAGTCTCTTCCATATATAAAGGCTTTCAGTTAACTCACTGAGG	1260		361	GAGGAGTGTGCTGTGTTTATCAAGGATCCACAAATGCTAAAGAGCTGTCTTCCAAAGGA	420	
Db	1201	GGCAAAATGTAATGTCGAGTCTCTTCCATATATAAAGGCTTTCAGTTAACTCACTGAGG	1260		361	GAGGAGTGTGCTGTGTTTATCAAGGATCCACAAATGCTAAAGAGCTGTCTTCCAAAGGA	420	
Qy	1261	GTATCTGCAATGCTGAGTGTATGAACAAAGTGAGGAGATGAATGTATGCTCTTAG	1320		421	GTCAAAATCTGGGATGCCAATGATCCGAGACTTTTGGACAGCTTGGGATTTCTCCACC	480	
Db	1261	GTATCTGCAATGCTGAGTGTATGAACAAAGTGAGGAGATGAATGTATGCTCTTAG	1320		421	GTCAAAATCTGGGATGCCAATGATCCGAGACTTTTGGACAGCTTGGGATTTCTCCACC	480	
Qy	1321	CAAAACATGATGTGCAATTTCAATCCAGTCTATATAAAGAGGTTGGTGAATTTTAC	1380		481	AGAGAAAGAGGGGACTTGGGCCCCAGTTTATGGCTTCCAGTGGAGGCAATTTGGGGCAGAA	540	
Db	1321	CAAAACATGATGTGCAATTTCAATCCAGTCTATATAAAGAGGTTGGTGAATTTTAC	1380		481	AGAGAAAGAGGGGACTTGGGCCCCAGTTTATGGCTTCCAGTGGAGGCAATTTGGGGCAGAA	540	
Qy	1381	AGCTATTTTGGAAATATTTTAGAATTTTAAAGATTTTCAAGCTATTTCCCTCAAAAT	1440		541	TACAGAGATPATGAATCAGATTATTCAGGACAGGAGTGTGACCACTGCAAGAGTGATT	600	
Db	1381	AGCTATTTTGGAAATATTTTAGAATTTTAAAGATTTTCAAGCTATTTCCCTCAAAAT	1440		541	TACAGAGATPATGAATCAGATTATTCAGGACAGGAGTGTGACCACTGCAAGAGTGATT	600	
Qy	1441	CTGAGGAGCTGAGTAACACCATCGATCATGATGATGAGTGTGTTATGAATTTTANAGT	1500		601	GACACCATCAAAACCAACCTTGACGACAGAGAAATCATCATGTGGCTTGGAAATCCAAGA	660	
Db	1441	CTGAGGAGCTGAGTAACACCATCGATCATGATGATGAGTGTGTTATGAATTTTANAGT	1500		601	GACACCATCAAAACCAACCTTGACGACAGAGAAATCATCATGTGGCTTGGAAATCCAAGA	660	
Qy	1501	TGTTTTATATGTTGCTATATAAAGAGTGTCTTGC	1536		721	AGTGAGCTGTCTGCGCAGCTGTACAGAGATCGGAGACATGGGCTCGGTGTGCTTTC	780	
Db	1501	TGTTTTATATGTTGCTATATAAAGAGTGTCTTGC	1536		721	AGTGAGCTGTCTGCGCAGCTGTACAGAGATCGGAGACATGGGCTCGGTGTGCTTTC	780	
RESULT 3								
US-09-442-384A-792								
; Sequence 792, Application US/09442384A								
; GENERAL INFORMATION:								
; APPLICANT: Chenchik, Alex								
; APPLICANT: Lukashov, Matvey								
; TITLE OF INVENTION: Hematology/Immunology Array								
; FILE REFERENCE: CLON-006CIP15								
; CURRENT APPLICATION NUMBER: US/09/442,384A								
; CURRENT FILING DATE: 1999-11-17								
; PRIOR APPLICATION NUMBER: 09/053,375								
; PRIOR FILING DATE: 1998-03-31								
; NUMBER OF SEQ ID NOS: 830								
; SOFTWARE: FastSeq for Windows Version 4.0								
; SEQ ID NO 792								
; LENGTH: 1536								
; TYPE: DNA								
; ORGANISM: homo sapiens								
US-09-442-384A-792								
Query Match 99.8%; Score 1533; DB 18; Length 1536;								
Best Local Similarity 99.8%; Pred. No. 0;								
Matches 1533; Conservative 0; Mismatches 3; Indels 0; Gaps 0;								
Qy	1	GGGGGGGGGGGACACATTTGGCTTCCCTCCGCTCCCGCCCGCCCATTTGGCTCCCTCGT	60		1081	GTCTTTAGGGTGGGCTGGATGCCGAGGTAAAGTCTCTTTTCCTCTAAAGAAAAGG	1140	
Db	1	GGGGGGGGGGGACACATTTGGCTTCCCTCCGCTCCCGCCCGCCCATTTGGCTCCCTCGT	60		1081	GTCTTTAGGGTGGGCTGGATGCCGAGGTAAAGTCTCTTTTCCTCTAAAGAAAAGG	1140	
Qy	61	CCCGCCCGCCCATTTGGCTTCCCTCCGCTCCCGCCCGCCCATTTGGCTTGGCCGGC	120		1141	AACCTAGTCAAAATCTGTCGGTCACTACGATTAATTAATTTTAAAGGATGTGCCACT	1200	
Db	61	CCCGCCCGCCCATTTGGCTTCCCTCCGCTCCCGCCCGCCCATTTGGCTTGGCCGGC	120		1141	AACCTAGTCAAAATCTGTCGGTCACTACGATTAATTAATTTTAAAGGATGTGCCACT	1200	
Qy	121	TGGAGTGTCCGCGCGGCTTGGCTTCCCTCCGCTCCCGCCCGCCCATTTGGCTTGGCCGGC	180		1201	GGCAAAATGTAATGTCGAGTCTCTTCCATATATAAAGGCTTTCAGTTAACTCACTGAGG	1260	
Db	121	TGGAGTGTCCGCGCGGCTTGGCTTCCCTCCGCTCCCGCCCGCCCATTTGGCTTGGCCGGC	180		1201	GGCAAAATGTAATGTCGAGTCTCTTCCATATATAAAGGCTTTCAGTTAACTCACTGAGG	1260	
Qy	181	CCGCGGACGGGAGCTGAGTACCTGGGGAGATCCAACATCTCCGCTGGCGGCTC	240		1261	GTATCTGCAATGCTGAGTGTATGAACAAAGTGAGGAGATGAATGTATGCTCTTAG	1320	
Db	181	CCGCGGACGGGAGCTGAGTACCTGGGGAGATCCAACATCTCCGCTGGCGGCTC	240		1261	GTATCTGCAATGCTGAGTGTATGAACAAAGTGAGGAGATGAATGTATGCTCTTAG	1320	
Qy	241	AGGAGGACGACCGCACGGGACCGGCACTCTGTGGTATTCGGCATCGAGCGCGCTAC	300		1321	CAAAACATGATGTGCAATTTCAATCCAGTCTATATAAAGAGGTTGGTGAATTTTAC	1380	

Db	1321	CAAAAACATGATGTGTCGCAATTCAAATCCCAACGTAATTAAAGAGAGGTGTGTTGATTCAC	1380
Qy	1381	AAGCTATTTTGGGAATATTTTGAATAATTTTAAAGAAATTTCAAGCTATTCCTCTCAAT	1440
Db	1381	AAGCTATTTTGGGAATATTTTGAATAATTTTAAAGAAATTTCAAGCTATTCCTCTCAAT	1440
Qy	1441	CTGAGGAGCTCAGTAACACCATTCGATCATGTGTAGAGTGTGGTTATGAACTTTANAGT	1500
Db	1441	CTGAGGAGCTGAGTAACACCATTCGATCATGTGTAGAGTGTGGTTATGAACTTTATAGT	1500
Qy	1501	TGTTTTATATGTTGCTATAATAAAGAGTGTTCGC	1536
Db	1501	TGTTTTATATGTTGCTATAATAAAGAGTGTTCGC	1536

RESULT 4

US-09-442-384B-792

; Sequence 792, Application US/09442384B

; GENERAL INFORMATION:

; APPLICANT: Chenchik, Alex

; APPLICANT: Lukashiev, Marvey

; TITLE OF INVENTION: Hematology/Immunology Array

; FILE REFERENCE: CLON-006CIP15

; CURRENT APPLICATION NUMBER: US/09/442,384B

; PRIOR FILING DATE: 1999-11-17

; PRIOR APPLICATION NUMBER: 09/053,375

; PRIOR FILING DATE: 1998-03-31

; NUMBER OF SEQ ID NOS: 830

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 792

; LENGTH: 1536

; TYPE: DNA

; ORGANISM: homo sapiens

US-09-442-384B-792

Qy	1	GGGGGGGGGGGACACATTGGCTTGCTCCGTCGGCGGGGACATTGGCGTCGCTCCGT	60
Db	1	GGGGGGGGGGGAGACACATTGGCTTGCTCCGTCGGCGGGGACATTGGCGTCGCTCCGT	60
Qy	61	CCCGCGCGGCGCATTGCGCTCCGTCGTCGTCGCGCGCGCGCCATGCTGTGGCGCGC	120
Db	61	CCCGCGCGGCGCATTGCGCTCCGTCGTCGTCGCGCGCGCGCCATGCTGTGGCGCGC	120
Qy	121	TGCGAGCTGCGCGCGCGCGCTTCGCCCGCGCGCGCACAGGAGGGGAGCGCGAGCGCGT	180
Db	121	TGCGAGCTGCGCGCGCGCGCTTCGCCCGCGCGCGCACAGGAGGGGAGCGCGAGCGCGT	180
Qy	181	CGCGCGCACGGGAGCTGCAGTACTTGGGGGAGATCCAAACACATCTCCGTCGGCGGTC	240
Db	181	CGCGCGCACGGGAGCTGCAGTACTTGGGGGAGATCCAAACACATCTCCGTCGGCGGTC	240
Qy	241	AGGAAGGACGACCGCACGGGACCGGCAACCTGTGCGTATTCGGCATGCAAGCGCGCTAC	300
Db	241	AGGAAGGAGNACCGCACGGGACCGGCAACCTGTGCGTATTCGGCATGCAAGCGCGCTAC	300
Qy	301	AGCCTGAGAGAGAAATTCCTCTGTCGACACCAACCTGTGTCTGGAAGGAGGTTTTC	360
Db	301	AGCCTGAGAGAGAAATTCCTCTGTCGACACCAACCTGTGTCTGGAAGGAGGTTTTC	360
Qy	361	GAGGAGTGTGCTGTTTATCAAGGATCCACAAATGCTAAAGAGCTGTCTCCAGGGA	420
Db	361	GAGGAGTGTGCTGTTTATCAAGGATCCACAAATGCTAAAGAGCTGTCTCCAGGGA	420
Qy	421	GTGAAATCTGGAGTGCATGATGCCAGAGCTTTTGGAGACGCTTGGATTTCTCAC	480
Db	421	GTGAAATCTGGAGTGCATGATGCCAGAGCTTTTGGAGACGCTTGGATTTCTCAC	480
Qy	481	AGAGAGAGGGGACTTTGGGCCCGAGTTATGCGTCTCCAGTGGAGGATTTTGGGCGAG	540
Db	481	AGAGAGAGGGGACTTTGGGCCCGAGTTATGCGTCTCCAGTGGAGGATTTTGGGCGAG	540

Query Match 99.8%; Score 1533; DB 18; Length 1536;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1533; Conservative 0; Mismatches 3; Indels 0; Gaps

```

; Sequence 141, Application US/09631275
; GENERAL INFORMATION:
; APPLICANT: Hong Chen
; APPLICANT: Nelson B. Freimer
; APPLICANT: Thomas Novak
; TITLE OF INVENTION: Method and Compositions for Diagnosing and Treating Chromosome-18
; FILE REFERENCE: 7853-146-999
; CURRENT APPLICATION NUMBER: US/09/631,275
; CURRENT FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 09/268,992
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/236,134
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/078,044
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: 60/088,312
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/106,056
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 158
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 141
; LENGTH: 1536
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-631-275-141

Query Match          99.8%; Score 1533; DB 24; Length 1536;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGACACCTTGCGCTGCTCCGCTCCGCGCGACCTTGCGCTCCGCT 60
DB 1 GGGGGGGGGGGGGGACACCTTGCGCTGCTCCGCTCCGCGCGACCTTGCGCTCCGCT 60
QY 61 CCCCCCGCGCACCTTGCGCTGCTCCGCTCCGCGCGACCTTGCGCTCCGCT 120
DB 61 CCCCCCGCGCACCTTGCGCTGCTCCGCTCCGCGCGACCTTGCGCTCCGCT 120
QY 121 TGGAGAGTGCCTGCGCGCGCGCTTGCGCGCGCGACAGAGAGCGAGCGCGCT 180
DB 121 TGGAGAGTGCCTGCGCGCGCGCTTGCGCGCGCGACAGAGAGCGAGCGCGCT 180
QY 181 CCGCGCGCGCGCGAGCTGACATCTGCGCGCGAGATCCAAACATCTCCGCTGCGCGCT 240
DB 181 CCGCGCGCGCGCGAGCTGACATCTGCGCGCGAGATCCAAACATCTCCGCTGCGCGCT 240
QY 241 AGAAGAGAGAGCGCGACCGCGCGACCGCGACCTGCGGATTTGGGCGAGCGCGCTAC 300
DB 241 AGAAGAGAGAGCGCGACCGCGCGACCGCGACCTGCGGATTTGGGCGAGCGCGCTAC 300
QY 301 AGCTGAGAGATGAATTCCTCTGCTGACAAACAAAGTGTCTTGGAGAGGCTTTTG 360
DB 301 AGCTGAGAGATGAATTCCTCTGCTGACAAACAAAGTGTCTTGGAGAGGCTTTTG 360
QY 361 GAGAGAGTGTGCTGCTTATCAAGGGATCCAAATCTTAAAGAGTCTTCCAAAGGGA 420
DB 361 GAGAGAGTGTGCTGCTTATCAAGGGATCCAAATCTTAAAGAGTCTTCCAAAGGGA 420
QY 421 GTGAAAATCTGGGATGCCAATGATCCGAGACTTTTGGAGAGCTTGGGATTCCTCAAC 480
DB 421 GTGAAAATCTGGGATGCCAATGATCCGAGACTTTTGGAGAGCTTGGGATTCCTCAAC 480
QY 481 AGAGAGAGAGGAGCTTGGGCGCACTTATGCTTCCAGTGAAGGCACTTTTGGGCGAGAA 540
DB 481 AGAGAGAGAGGAGCTTGGGCGCACTTATGCTTCCAGTGAAGGCACTTTTGGGCGAGAA 540
QY 541 TACAGAGATGTGAATCAATATTTTCAAGAGAGGAGATTTCACTGCAAGAGAGTAT 600
DB 541 TACAGAGATGTGAATCAATATTTTCAAGAGAGGAGATTTCACTGCAAGAGAGTAT 600
QY 601 GACACGATCAAAACCACTTACGACAGAGAGATATATATGCGCTTGGAGATCAAGA 660

```

```

DB 601 GACACGATCAAAACCACTTACGACAGAGAGATATATATGCGCTTGGAGATCAAGA 660
QY 661 GATCTCTCTGATGAGCGCTGCTCCATGCGATGCGCTCCGAGTCTATGAGTGAAC 720
DB 661 GATCTCTCTGATGAGCGCTGCTCCATGCGATGCGCTCCGAGTCTATGAGTGAAC 720
QY 721 AGTACGCTGCTGCGACGCTGACAGAGATGCGGAGACATGAGGCTCGGAGGCTTTC 780
DB 721 AGTACGCTGCTGCGACGCTGACAGAGATGCGGAGACATGAGGCTCGGAGGCTTTC 780
QY 781 AAGATGCGAGGATGAGCGCTGCTGACATGATGAGTGGCAATCAAGCGGCTGAAGCCA 840
DB 781 AAGATGCGAGGATGAGCGCTGCTGACATGATGAGTGGCAATCAAGCGGCTGAAGCCA 840
QY 841 GGTGACCTTATACACTTGGAGATGACATATTTTACCTGATATCAATGAGGACCTG 900
DB 841 GGTGACCTTATACACTTGGAGATGACATATTTTACCTGATATCAATGAGGACCTG 900
QY 901 AAAATTCAGCTTACGAGAAACCAAGCTTTCCAAAGCTGAGGATTTCCGAAAAGTT 960
DB 901 AAAATTCAGCTTACGAGAAACCAAGCTTTCCAAAGCTGAGGATTTCCGAAAAGTT 960
QY 961 GAGAAATTTGATGACTTCAAGCTGAAGACTTTCAGATTGAAGGATCAATCCGATCCA 1020
DB 961 GAGAAATTTGATGACTTCAAGCTGAAGACTTTCAGATTGAAGGATCAATCCGATCCA 1020
QY 1021 ACTATTTAAATGGAATGCGCTGTTAGGCTTCAAGAGAGCTGAGATATTTCTCA 1080
DB 1021 ACTATTTAAATGGAATGCGCTGTTAGGCTTCAAGAGAGCTGAGATATTTCTCA 1080
QY 1081 GCTTTTGGGGTGGGCTGAGATGCGGAGTAAAGTCTTTTGTCTTAAAGAAAGG 1140
DB 1081 GCTTTTGGGGTGGGCTGAGATGCGGAGTAAAGTCTTTTGTCTTAAAGAAAGG 1140
QY 1141 AACTAGGTCAAAATCTGCGGTGACCTATGATATTTTAAAGATGTGCGACT 1200
DB 1141 AACTAGGTCAAAATCTGCGGTGACCTATGATATTTTAAAGATGTGCGACT 1200
QY 1201 GCGAAATGATCTGCGGCTTCTTCCATATTAAGGCTTGAATCACTGAGG 1260
DB 1201 GCGAAATGATCTGCGGCTTCTTCCATATTAAGGCTTGAATCACTGAGG 1260
QY 1261 GATCTGACATGAGGTTATGAACAAAGTGAAGATGAATGATGCTTTAG 1320
DB 1261 GATCTGACATGAGGTTATGAACAAAGTGAAGATGAATGATGCTTTAG 1320
QY 1321 CAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
DB 1321 CAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 1381 AAGCTATTTTGGAAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 1440
DB 1381 AAGCTATTTTGGAAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 1440
QY 1441 CTGAGAGAGCTGAGTAAACATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1441 CTGAGAGAGCTGAGTAAACATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1501 TGTATTTATGCTGATATTAAGAGTGTTCG 1536
DB 1501 TGTATTTATGCTGATATTAAGAGTGTTCG 1536

RESULT 6
US-09-638-659-6
; Sequence 6, Application US/09658659
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: TO POLAR METABOLISM HAVING UTILITY IN DETERMINING THE
; FILE REFERENCE: 11926-015001

```

```

; CURRENT APPLICATION NUMBER: US/09/658,659
; CURRENT FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1536
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1066
; OTHER INFORMATION: n = t or c
; NAME/KEY: misc_feature
; LOCATION: 1136
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc_feature
; LOCATION: 1497
; OTHER INFORMATION: n = t or a
; US-09-658-659-6

```

```

Query Match          99.8%; Score 1533; DB 25; Length 1536;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGGGGGGGGGAGCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGCT 60
DB 1 GGGGGGGGGGAGCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGCT 60
QY 61 CCGCGCGCGCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGCT 120
DB 61 CCGCGCGCGCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGCT 120
QY 121 TCGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 180
DB 121 TCGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 180
QY 181 CCGCGCGCGAGCTGAGTCACTGCGCGCGAGTCACTGCGCGCGCT 240
DB 181 CCGCGCGCGAGCTGAGTCACTGCGCGCGAGTCACTGCGCGCGCT 240
QY 241 AGAAGGACGACCGGACCGGAGCCGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 241 AGAAGGACGACCGGAGCGGAGCCGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 301 AGCTGAGAGTGAATTCCTCTGCTGAGCAACCAACCTGCTGCTGCTGCT 360
DB 301 AGCTGAGAGTGAATTCCTCTGCTGAGCAACCAACCTGCTGCTGCTGCT 360
QY 361 GAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 361 GAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 421 GTGAAATCTGGAGTCCGATGATCCGAGACTTTTGGACAGCTGGGATTT 480
DB 421 GTGAAATCTGGAGTCCGATGATCCGAGACTTTTGGACAGCTGGGATTT 480
QY 481 AGAAGAGAGGAGCTTGGGCGCGCTTATGCTTCAATGAGAGGATTTTGG 540
DB 481 AGAAGAGAGGAGCTTGGGCGCGCTTATGCTTCAATGAGAGGATTTTGG 540
QY 541 TACAGAGATATGAGATCAAGATTTATTCAGGACAGGAGGCTTACCAAGAT 600
DB 541 TACAGAGATATGAGATCAAGATTTATTCAGGACAGGAGGCTTACCAAGAT 600
QY 601 GACACCATCAAAACCACTGACGACAGAGAGATCAGTGGCTTGAATCCAGA 660

```

```

DB 601 GACACCATCAAAACCACTGACGACAGAGAGATCAGTGGCTTGAATCCAGA 660
QY 661 GATCTTCCTGAGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 GATCTTCCTGAGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 721 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 781 AACATGCGACGCTAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 781 AACATGCGACGCTAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 841 GGTGACTTTATACACTTTGGAGATGCAATATTTACTGATGACATCGAG 900
DB 841 GGTGACTTTATACACTTTGGAGATGCAATATTTACTGATGACATCGAG 900
QY 901 AAAATTCAGCTTCAGGAGAACCCAGACCTTCCGAAAGCTGAGATTCGAA 960
DB 901 AAAATTCAGCTTCAGGAGAACCCAGACCTTCCGAAAGCTGAGATTCGAA 960
QY 961 GAGAAATTTGATGACTTCAAGCTGAGACTTCAAGTGAAGGATCAATCCGAT 1020
DB 961 GAGAAATTTGATGACTTCAAGCTGAGACTTCAAGTGAAGGATCAATCCGAT 1020
QY 1021 ACTATTAATATGAAATGAGCTGTTAGGCTTCAAGAGCTGAGATATTTGCA 1080
DB 1021 ACTATTAATATGAAATGAGCTGTTAGGCTTCAAGAGCTGAGATATTTGCA 1080
QY 1081 GTCTTTAGGGGCTGAGTGCAGGATGCGAGGTAAGTCTTTTGTCTTAAAG 1140
DB 1081 GTCTTTAGGGGCTGAGTGCAGGATGCGAGGTAAGTCTTTTGTCTTAAAG 1140
QY 1141 AACTAGGTCAAAATCTGCGCGGCTATCATGATTAATTTTAAAGATGCGACT 1200
DB 1141 AACTAGGTCAAAATCTGCGCGGCTATCATGATTAATTTTAAAGATGCGACT 1200
QY 1201 GGCAAATGAACTGCGCACTTCTTCAATATAAAGCTTTGAGTTACTCATGAG 1260
DB 1201 GGCAAATGAACTGCGCACTTCTTCAATATAAAGCTTTGAGTTACTCATGAG 1260
QY 1261 GTATCTGCAATGCTGAGTATGAAACAAAGTGAAGGAAATGATGCTCTT 1320
DB 1261 GTATCTGCAATGCTGAGTATGAAACAAAGTGAAGGAAATGATGCTCTT 1320
QY 1321 CAAAACATGATGATGATTTCAATCCAGTACTTAAAGAGGTGGGAAATTCAC 1380
DB 1321 CAAAACATGATGATGATTTCAATCCAGTACTTAAAGAGGTGGGAAATTCAC 1380
QY 1381 AAGCTATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAAT 1440
DB 1381 AAGCTATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAAT 1440
QY 1441 CTGAGGAGCTGAGTAAACCAATCGATCATGATGATGATGATGATGAT 1500
DB 1441 CTGAGGAGCTGAGTAAACCAATCGATCATGATGATGATGATGATGATGAT 1500
QY 1501 TGTTTTATGCTGCTATATTAAGAGTGTCTGCT 1536
DB 1501 TGTTTTATGCTGCTATATTAAGAGTGTCTGCT 1536

```

```

RESULT 7
US-09-877-178-11
; Sequence 11, Application US/09877178
; GENERAL INFORMATION:
; APPLICANT: Kathleen Danenberg, et al.
; TITLE OF INVENTION: METHOD OF DETERMINING A CHEMOTHERAPEUTIC
; TITLE OF INVENTION: REGIMEN BASED ON REC1 and TS EXPRESSION
; FILE REFERENCE: 11220-119
; CURRENT APPLICATION NUMBER: US/09/877,178

```

[illegible]

```

1 RESULT 8
2 US-09-962-665-6
3 / Sequence 6, Application US/09962665
4 /
5 / GENERAL INFORMATION:
6 / APPLICANT: Statton, Jr., Vincent P.
7 / TITLE OF INVENTION: POLY(DIOLYDUMATE SYNTHETASE GENE SEQUENCE
8 / TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
9 / TITLE OF INVENTION: TREATMENT OF DISEASE
10 / FILE REFERENCE: 11926-015004
11 /
12 / CURRENT APPLICATION NUMBER: US/09/962,665
13 / CURRENT FILING DATE: 2001-09-24
14 /
15 / PRIOR APPLICATION NUMBER: 09/658,659
16 / PRIOR FILING DATE: 2000-09-08
17 /
18 / PRIOR APPLICATION NUMBER: 09/556,033
19 / PRIOR FILING DATE: 2000-06-15
20 /
21 / PRIOR APPLICATION NUMBER: 09/357,743
22 / PRIOR FILING DATE: 1999-07-20
23 /
24 / PRIOR APPLICATION NUMBER: 09/357,024
25 / PRIOR FILING DATE: 1999-07-19
26 /
27 / PRIOR APPLICATION NUMBER: 60/093,484
28 / PRIOR FILING DATE: 1998-07-20
29 /
30 / NUMBER OF SEQ ID NOS: 16
31 /
32 / SOFTWARE: FastSeq for Windows Version 4.0
33 /
34 / SEQ ID NO 6
35 / LENGTH: 1536
36 /
37 / TYPE: DNA
38 /
39 / ORGANISM: Homo sapiens
40 /
41 / FEATURE:

```

NAME/KEY: misc feature
 LOCATION: 1066
 OTHER INFORMATION: n = t or c
 NAME/KEY: misc feature
 LOCATION: 1136
 OTHER INFORMATION: n = a or g
 NAME/KEY: misc feature
 LOCATION: 1497
 OTHER INFORMATION: n = t or a
 US-09-962-665-6

Query Match 99.8%; Score 1533; DB 36; Length 1536;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGACCTTGGGAGATGACATATTACCTGATCATGACCACTG 60
 Db 1 GGGGGGGGGGGGACCTTGGGAGATGACATATTACCTGATCATGACCACTG 60
 Qy 61 CCGGCGGCGGACCTTGGGAGATGACATATTACCTGATCATGACCACTG 120
 Db 61 CCGGCGGCGGACCTTGGGAGATGACATATTACCTGATCATGACCACTG 120
 Qy 121 TCGGAGCTGCGGCGGCGGACCTTGGGAGATGACATATTACCTGATCATG 180
 Db 121 TCGGAGCTGCGGCGGCGGACCTTGGGAGATGACATATTACCTGATCATG 180
 Qy 181 CCGGCGGCGGAGCTGACATGACCTTGGGAGATGACATATTACCTGATCATG 240
 Db 181 CCGGCGGCGGAGCTGACATGACCTTGGGAGATGACATATTACCTGATCATG 240
 Qy 241 AGGAGAGACGACCGGAGCTGACATGACCTTGGGAGATGACATATTACCTG 300
 Db 241 AGGAGAGACGACCGGAGCTGACATGACCTTGGGAGATGACATATTACCTG 300
 Qy 301 AGCTGAGAGATGATTCCTGCTGCTGACCAACAGTGTGCTTGGAGAGGCTT 360
 Db 301 AGCTGAGAGATGATTCCTGCTGCTGACCAACAGTGTGCTTGGAGAGGCTT 360
 Qy 361 GAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Db 361 GAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Qy 421 GTGAAATCTGGAGATGACATGACCTTGGGAGATGACATATTACCTGATCATG 480
 Db 421 GTGAAATCTGGAGATGACATGACCTTGGGAGATGACATATTACCTGATCATG 480
 Qy 481 AGGAGAGAGGAGCTTGGGAGATGACATATTACCTGATCATGACCACTT 540
 Db 481 AGGAGAGAGGAGCTTGGGAGATGACATATTACCTGATCATGACCACTT 540
 Qy 541 TACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 Db 541 TACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 Qy 601 GACACATCAAAACCAACCTGACGAGAGAGATGATGATGATGATGATGAT 660
 Db 601 GACACATCAAAACCAACCTGACGAGAGAGATGATGATGATGATGATGAT 660
 Qy 661 GATCTTCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Db 661 GATCTTCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Qy 721 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Db 721 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Qy 781 AACATGCGACGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 781 AACATGCGACGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Qy 841 GGTGACTTTATACACTTTGGAGATGACATATTACCTGATCATGACCACTG 900

Db 841 GGTGACTTTATACACTTTGGAGATGACATATTACCTGATCATGACCACTG 900
 Qy 901 AAAATTCAGCTTGGGAGACCTGACATCTTCCCAAGCTGAGATTTCTGAAAATT 960
 Db 901 AAAATTCAGCTTGGGAGACCTGACATCTTCCCAAGCTGAGATTTCTGAAAATT 960
 Qy 961 GAGAAATTCAGCTTGGGAGACCTGACATCTTCCCAAGCTGAGATTTCTGAAAATT 1020
 Db 961 GAGAAATTCAGCTTGGGAGACCTGACATCTTCCCAAGCTGAGATTTCTGAAAATT 1020
 Qy 1021 ACTATTTAAATGAGATGAGCTGTTTAAAGCTTCCCAAGCTGAGATTTCTGAAAATT 1080
 Db 1021 ACTATTTAAATGAGATGAGCTGTTTAAAGCTTCCCAAGCTGAGATTTCTGAAAATT 1080
 Qy 1081 GTCTTTAGGGGCTTGGGAGATGAGCTGTTTAAAGCTTCCCAAGCTGAGATTTCTGAAAATT 1140
 Db 1081 GTCTTTAGGGGCTTGGGAGATGAGCTGTTTAAAGCTTCCCAAGCTGAGATTTCTGAAAATT 1140
 Qy 1141 AACTAGTCAAAATGAGCTGTTTAAAGCTTCCCAAGCTGAGATTTCTGAAAATT 1200
 Db 1141 AACTAGTCAAAATGAGCTGTTTAAAGCTTCCCAAGCTGAGATTTCTGAAAATT 1200
 Qy 1201 GGCATATGATGAGCTGTTTAAAGCTTCCCAAGCTGAGATTTCTGAAAATT 1260
 Db 1201 GGCATATGATGAGCTGTTTAAAGCTTCCCAAGCTGAGATTTCTGAAAATT 1260
 Qy 1261 GTATCTGACATGCTGAGTTTAAAGCTTCCCAAGCTGAGATTTCTGAAAATT 1320
 Db 1261 GTATCTGACATGCTGAGTTTAAAGCTTCCCAAGCTGAGATTTCTGAAAATT 1320
 Qy 1321 CAAAATGATGAGCTTCCCAAGCTTCCCAAGCTTCCCAAGCTTCCCAAGCTT 1380
 Db 1321 CAAAATGATGAGCTTCCCAAGCTTCCCAAGCTTCCCAAGCTTCCCAAGCTT 1380
 Qy 1381 AAGCTATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 1440
 Db 1381 AAGCTATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 1440
 Qy 1441 CTGAGGAGCTGAGTAAACCATCATGATGATGATGATGATGATGATGATGAT 1500
 Db 1441 CTGAGGAGCTGAGTAAACCATCATGATGATGATGATGATGATGATGATGAT 1500
 Qy 1501 TGTATTATGTTCTATATTAAGAAGTGTCTGC 1536
 Db 1501 TGTATTATGTTCTATATTAAGAAGTGTCTGC 1536

RESULT 9
 US-09-962-677-6
 Sequence 6, Application US/09962677
 GENERAL INFORMATION:
 APPLICANT: Stanton, Jr., Vincent P.
 TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
 TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
 FILE REFERENCE: 11926-015003
 CURRENT APPLICATION NUMBER: US/09/962,677
 PRIOR FILING DATE: 2001-09-24
 PRIOR APPLICATION NUMBER: 09/658,659
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 09/596,033
 PRIOR FILING DATE: 2000-06-15
 PRIOR APPLICATION NUMBER: 09/357,743
 PRIOR FILING DATE: 1999-07-20
 PRIOR APPLICATION NUMBER: 09/357,024
 PRIOR FILING DATE: 1999-07-19
 PRIOR APPLICATION NUMBER: 60/093,484
 PRIOR FILING DATE: 1998-07-20
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 6
 LENGTH: 1536
 TYPE: DNA


```

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1066
OTHER INFORMATION: n = t or c
NAME/KEY: misc_feature
LOCATION: 1136
OTHER INFORMATION: n = a or g
NAME/KEY: misc_feature
LOCATION: 1497
OTHER INFORMATION: n = t or a
US-09-962-677-6

```

```

Query Match          99.8%; Score 1533; DB 36; Length 1536;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGGGGGGGGGGGACCACTTGGCTCCGTCGCCGCGCCGACCTTGGCTCCGCT 60
DB 1 GGGGGGGGGGGGACCACTTGGCTCCGTCGCCGCGCCGACCTTGGCTCCGCT 60
QY 61 CCCGCGCGCACTTGGCTCCGTCGCCGCGCCGCGCCGCGCCGCGCCGCGCT 120
DB 61 CCCGCGCGCACTTGGCTCCGTCGCCGCGCCGCGCCGCGCCGCGCCGCGCT 120
QY 121 TCGGAGCTGCGCGCGCGCGCTTGGCTCCGTCGCCGCGCCGCGCCGCGCT 180
DB 121 TCGGAGCTGCGCGCGCGCGCTTGGCTCCGTCGCCGCGCCGCGCCGCGCT 180
QY 181 CCCGCGCGCACTTGGCTCCGTCGCCGCGCCGCGCCGCGCCGCGCT 240
DB 181 CCCGCGCGCACTTGGCTCCGTCGCCGCGCCGCGCCGCGCCGCGCT 240
QY 241 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 241 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 AGCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 301 AGCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 361 GAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 361 GAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 421 GTGAAATCTGGGATGCGCATGATCCGAGACTTTTGGACGCTTGGATCTCACC 480
DB 421 GTGAAATCTGGGATGCGCATGATCCGAGACTTTTGGACGCTTGGATCTCACC 480
QY 481 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 481 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 TACAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 600
DB 541 TACAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 600
QY 601 GACACATCAAAACCAACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 601 GACACATCAAAACCAACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 GATCTTCTCTGATGCGCTGCTTCAAGCAGATCCCTTGGCAGATTCATGCTG 720
DB 661 GATCTTCTCTGATGCGCTGCTTCAAGCAGATCCCTTGGCAGATTCATGCTG 720
QY 721 AGTAGAGCTGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 721 AGTAGAGCTGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 AACATGCGCAGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 781 AACATGCGCAGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840

```

```

QY 841 GGTGACTTTATACACTTTGGAGATGACATATTTACCTGATTCACATGAGCACTG 900
DB 841 GGTGACTTTATACACTTTGGAGATGACATATTTACCTGATTCACATGAGCACTG 900
QY 901 AAAATTCAGCTTCCAGAGAACCAACCTTCCCAAGCTCAGAGATCTTCCAAAAGTT 960
DB 901 AAAATTCAGCTTCCAGAGAACCAACCTTCCCAAGCTCAGAGATCTTCCAAAAGTT 960
QY 961 GAGAAATGATGACTTCAAGCTGAAGACTTTCAGATTTGAAGGATCAATCCGATCCA 1020
DB 961 GAGAAATGATGACTTCAAGCTGAAGACTTTCAGATTTGAAGGATCAATCCGATCCA 1020
QY 1021 ACTATTAATGAAATGCTGTTAGGCTTTCAAAAGAGCTTGAAGATTTGTGA 1080
DB 1021 ACTATTAATGAAATGCTGTTAGGCTTTCAAAAGAGCTTGAAGATTTGTGA 1080
QY 1081 GTCTTTAGGGGTTGGCTGAGATGCGAGGTTAAAGTTCTTTTGTCTTAAAGAAAGG 1140
DB 1081 GTCTTTAGGGGTTGGCTGAGATGCGAGGTTAAAGTTCTTTTGTCTTAAAGAAAGG 1140
QY 1141 AACTAGGTCAAAAATCTGCTCCGATCACTATCACTTATATTTTAAAGATGTCACCT 1200
DB 1141 AACTAGGTCAAAAATCTGCTCCGATCACTATCACTTATATTTTAAAGATGTCACCT 1200
QY 1201 GGCATATGTAATCTGCGAGTCTTTCATATPAAAGCTTGAATTACTCACTAGG 1260
DB 1201 GGCATATGTAATCTGCGAGTCTTTCATATPAAAGCTTGAATTACTCACTAGG 1260
QY 1261 GTATCTGACATCTGAGGTTATGACAAAGTGAAGAGATGAATGATGCTTGA 1320
DB 1261 GTATCTGACATCTGAGGTTATGACAAAGTGAAGAGATGAATGATGCTTGA 1320
QY 1321 CAATAATGATATGATCACTTCAATCCAGGTACTTAAAGAGGTTGGAATTTAC 1380
DB 1321 CAATAATGATATGATCACTTCAATCCAGGTACTTAAAGAGGTTGGAATTTAC 1380
QY 1381 AAGCTATTTTGGATATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 1440
DB 1381 AAGCTATTTTGGATATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 1440
QY 1441 CTGAGAGAGCTGAGTACACATCATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1441 CTGAGAGAGCTGAGTACACATCATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1501 TGTTTTATGTTGCTATATTAAGATGTTCTGC 1536
DB 1501 TGTTTTATGTTGCTATATTAAGATGTTCTGC 1536

```

```

RESULT 10
US-09-963-333-6
; Sequence 6, Application US/0996333
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES
; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT
; FILE REFERENCE: 11926-015002
; CURRENT APPLICATION NUMBER: US/09/963,333
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6

```



```

RESULT 11
US-10-172-118-556
; Sequence 556, Application US/10172118
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2639
; SEQ ID NO 556
; LENGTH: 1536
; TYPE: DNA

```

; ORGANISM: Homo sapiens
 ;
 ; PUBLICATION INFORMATION:
 ;
 ; DATABASE ACCESSION NUMBER: NM_0010711
 ;
 ; DATABASE ENTRY DATE: 2001-06-18
 ;
 US-10-172-118-556

Query Match	99.8%;	Score 1533;	DB 41;	Length 1536;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 1533; conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

[illegible]

QY	961	GAGAAATGTAAGACTTCGAAGCTGAGACCTTGAGTTGAAGGGTGAACATCCGACATCA	1020
Db	961	GAGAAATGTAAGACTTCGAAGCTGAGACCTTCAGATTGAAGGGTGAACATCCGACATCA	1020
QY	1021	ACTATTAAATGGAAATGGCGTTTGAAGGGTGTCTTCAAGGAGACNMGAGATATTGTCA	1080
Db	1021	ACTATTAAATGGAAATGGCGTTTGAAGGGTGTCTTCAAGGAGACCTTGAAGATATTGTCA	1080
QY	1081	GGCTTTGAGGGTTGGGCTCGAATGCGAAGGTAAAGTCTTTTGCTCTTAAAGANNAAG	1140
Db	1081	GGCTTTGAGGGTTGGGCTCGAATGCGAAGGTAAAGTCTTTTGCTCTTAAAGANNAAG	1140
QY	1141	AACTAAGGTCAAAAATCTGTCCGATGACCTTCACTAGTATTATTTTAAAGATGTGCACT	1200
Db	1141	AACTAAGGTCAAAAATCTGTCCGATGACCTTCACTAGTATTATTTTAAAGATGTGCACT	1200
QY	1201	GGCAATGTAACTGTGCAAGTTCCTTCCATTAATAAAGGCTTTGAGTTAACTCACTGAGG	1260
Db	1201	GGCAATGTAACTGTGCAAGTTCCTTCCATTAATAAAGGCTTTGAGTTAACTCACTGAGG	1260
QY	1261	GTATCTGACATCTGAGGTTATGACAAAGTGAGGAGATGAAATGTATGTGCTTTG	1320
Db	1261	GTATCTGACATCTGAGGTTATGACAAAGTGAGGAGATGAAATGTATGTGCTTTG	1320
QY	1321	CAAAAACATGATGTGCAATTCATCCGACGTACTTATAAAGAGGTTGGATTTTCAC	1380
Db	1321	CAAAAACATGATGTGCAATTCATCCGACGTACTTATAAAGAGGTTGGATTTTCAC	1380
QY	1381	AAGCTATTTTGGAAATATTTTGAANATTTTAAAGATTTCAACAATTCCTCCAAAT	1440
Db	1381	AAGCTATTTTGGAAATATTTTGAANATTTTAAAGATTTTCAACAATTCCTCCAAAT	1440
QY	1441	CTGAGGAGCTGTGATCAACATGATCATGATGATGATGAGGTGCTGTTATGAACTTAAAGT	1500
Db	1441	CTGAGGAGCTGTGATCAACATGATCATGATGATGATGAGGTGCTGTTATGAACTTAAAGT	1500
QY	1501	TGTTTTATATGTTGCTAATATTAAGAGATGTTCTGC	1536
Db	1501	TGTTTTATATGTTGCTAATATTAAGAGATGTTCTGC	1536

```

RESULT 12
US-60-412-049-350
Sequence 350, Application US/60412049
GENERAL INFORMATION:
APPLICANT: Baker, Joffe B.
APPLICANT: Cronin, Maureen T.
APPLICANT: Kiefer, Michael C.
APPLICANT: Shak, Steve
APPLICANT: Walker, Michael Graham
TITLE OF INVENTION: GENE EXPRESSION PROFILING IN BIOSIED TUMOR TISSUES
FILE REFERENCE: GENHE.004PR
CURRENT APPLICATION NUMBER: US/60/412,049
CURRENT FILING DATE: 2002-09-18
NUMBER OF SEQ ID NOS: 384
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 350
LENGTH: 1536
TYPE: DNA
ORGANISM: Homo sapiens
US-60-412-049-350

Query Match          99.8%; Score 1533; DB 85; Length 1536;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 3; Indels 0; Gaps

QY      1 GGGGGGGGGGGGACATTGGGCTCGCTCCGCGCGCGCACATTGGGCTCGCTCGGT 60
Db      1 GGGGGGGGGGGGACATTGGGCTCGCTCCGTCGCGCGCGCGCGCACATTGGGCTCGGT 60
QY      61 CCGCGCGCGCACATTGGGCTCGCTCCGTCGCGCGCGCGCGCGCGCATGCTGTGGCGCGGC 1200

```


311 ATGATTCCTCTGCTGACAAACCAAGTGTTCGGAAGAGGCTGTTTGGAGAGTTC 370
345 ATGAATTCCTCTGCTGACAAACCAAGTGTTCGGAAGAGGCTGTTTGGAGAGTTC 404
371 TGTGGTTTATCAAGAGATCCCAAAATGCTAAAGAGCTGCTTCCAGAGAGTGAATCT 430
405 TGTGGTTTATCAAGAGATCCCAAAATGCTAAAGAGCTGCTTCCAGAGAGTGAATCT 464
431 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGGAGATCTCCACAGAGAGAG 490
465 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGGAGATCTCCACAGAGAGAG 524
491 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGGAGATCTCCACAGAGAT 550
525 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGGAGATCTCCACAGAGAT 584
551 TGGATTCAGATTTATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 610
585 TGGATTCAGATTTATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 644
611 AAACCAACCTGACGACAGAGAGATATATATATATATATATATATATATATAT 670
645 AAACCAACCTGACGACAGAGAGATATATATATATATATATATATATATATAT 704
671 TGTATGCGCTGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 730
705 TGTATGCGCTGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 764
731 CCTGCGAGCTGACGACAGAGATATATATATATATATATATATATATATATAT 790
765 CCTGCGAGCTGACGACAGAGATATATATATATATATATATATATATATATAT 824
791 GCTATGCGCTGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 850
825 GCTATGCGCTGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 884
851 TACACACTTTGGAGAGATGACATATATATATATATATATATATATATATATAT 910
885 TACACACTTTGGAGAGATGACATATATATATATATATATATATATATATATAT 944
911 TACACACTTTGGAGAGATGACATATATATATATATATATATATATATATATAT 970
945 TACACACTTTGGAGAGATGACATATATATATATATATATATATATATATATAT 1004
971 ATGACTTCAAGCTGAGAGATTTTCAAGAGAGATTTTCAAGAGAGATTTTCAAG 1030
1005 ATGACTTCAAGCTGAGAGATTTTCAAGAGAGATTTTCAAGAGAGATTTTCAAG 1064
1031 TGAATATGCTGTTTGGAGAGATTTTCAAGAGAGATTTTCAAGAGAGATTTTCAAG 1090
1065 TGAATATGCTGTTTGGAGAGATTTTCAAGAGAGATTTTCAAGAGAGATTTTCAAG 1124
1091 GTTGGCTGAGATGCTGAGAGATTTTCAAGAGAGATTTTCAAGAGAGATTTTCAAG 1150
1125 GTTGGCTGAGATGCTGAGAGATTTTCAAGAGAGATTTTCAAGAGAGATTTTCAAG 1184
1151 AAAATCTGCTGAGAGATTTTCAAGAGAGATTTTCAAGAGAGATTTTCAAGAGAT 1210
1185 AAAATCTGCTGAGAGATTTTCAAGAGAGATTTTCAAGAGAGATTTTCAAGAGAT 1244
1211 ACTGTCAGATTTTCAAGAGATTTTCAAGAGAGATTTTCAAGAGAGATTTTCAAG 1270
1245 ACTGTCAGATTTTCAAGAGATTTTCAAGAGAGATTTTCAAGAGAGATTTTCAAG 1304
1271 ATGCTGAGATTTTCAAGAGATTTTCAAGAGAGATTTTCAAGAGAGATTTTCAAG 1330
1305 ATGCTGAGATTTTCAAGAGATTTTCAAGAGAGATTTTCAAGAGAGATTTTCAAG 1364
1331 TATGCTGAGATTTTCAAGAGATTTTCAAGAGAGATTTTCAAGAGAGATTTTCAAG 1390
1365 TATGCTGAGATTTTCAAGAGATTTTCAAGAGAGATTTTCAAGAGAGATTTTCAAG 1424

1391 TGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTT 1450
1425 TGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTT 1484
1451 TGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTT 1504
1485 TGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTT 1544
1505 TGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTT 1536
1545 TGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTT 1576

RESULT 15
US-09-652-121-6686
; Sequence 6686, Application US/09652121
; GENERAL INFORMATION:
; APPLICANT: Disefano, Peter
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1188-001
; CURRENT APPLICATION NUMBER: US/09/652,121
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,129
; NUMBER OF SEQ ID NOS: 7615
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6686
; LENGTH: 1829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-121-6686

Query Match 98.0%; Score 1505.4; DB 25; Length 1829;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1522; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

11 GAGCACTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 70
45 GAGCACTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 104
71 CACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 130
105 CACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 164
131 CGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 190
165 CGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 224
191 GGGAGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 250
225 GGGAGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 284
251 ACCGCAAGGCGACCGGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 310
285 ACCGCAAGGCGACCGGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 344
311 ATGAATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 370
345 ATGAATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 404
371 TGTGGTTTATCAAGAGATCCCAAAATGCTAAAGAGCTGCTTCCAGAGAGTGAATCT 430
405 TGTGGTTTATCAAGAGATCCCAAAATGCTAAAGAGCTGCTTCCAGAGAGTGAATCT 464
431 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGGAGATCTCCACAGAGAGAG 490
465 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGGAGATCTCCACAGAGAGAG 524
491 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGGAGATCTCCACAGAGAGAG 550
525 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGGAGATCTCCACAGAGAGAG 584

QY 551 TGAATCAGATTATTCAGAGACGGAGTTACCACTGGAAGAAGTATGACCATCA 610
| | | | |
Db 585 TGAATCAGATTATTCAGAGACGGAGTTACCACTGGAAGAAGTATGACCATCA 644
| | | | |
QY 611 AAACCAACCCCTGACGACGAGAAATCATGCTGCGCTTGGAAATCCAGAGATCTTCCCTC 670
| | | | |
Db 645 AAACCAACCCCTGACGACGAGAAATCATGCTGCGCTTGGAAATCCAGAGATCTTCCCTC 704
| | | | |
QY 671 TGAATGCGCTGCTCCATGACCATGCTCTGCGAGTTCATGATGATGAGACATGAGCTGT 730
| | | | |
Db 705 TGAATGCGCTGCTCCATGACCATGCTCTGCGAGTTCATGATGATGAGACATGAGCTGT 764
| | | | |
QY 731 CCGGCAAGCTGTACAGAGATCGGAGACATGCGGCTCGGTGTCCTTTCAATGCGCA 790
| | | | |
Db 765 CCGGCAAGCTGTACAGAGATCGGAGACATGCGGCTCGGTGTCCTTTCAATGCGCA 824
| | | | |
QY 791 GCTACGCTGCTGCTCACTGACATGATGCGGACATCAAGGCTGAAAGCAGGTCCTTA 850
| | | | |
Db 825 GCTACGCTGCTGCTCACTGACATGATGCGGACATCAAGGCTGAAAGCAGGTCCTTA 884
| | | | |
QY 851 TACACCTTTGGAGATGACATATTAAGTATGACATGACGACATGAAATTCAGC 910
| | | | |
Db 885 TACACCTTTGGAGATGACATATTAAGTATGACATGACGACATGAAATTCAGC 944
| | | | |
QY 911 TTCAGGAGAACCCAGACCTTTCCCAAGCTCAGATTCCTGAAAAGTTGAGAAAATTG 970
| | | | |
Db 945 TTCAGGAGAACCCAGACCTTTCCCAAGCTCAGATTCCTGAAAAGTTGAGAAAATTG 1004
| | | | |
QY 971 ATGACTTCAAAGCTGAGACCTTCAGATGAAAGGTACATCCGATCCAACTATTAAA 1030
| | | | |
Db 1005 ATGACTTCAAAGCTGAGACCTTCAGATGAAAGGTACATCCGATCCAACTATTAAA 1064
| | | | |
QY 1031 TGGAAATGCTGTTTAAAGGCTCTTCAAAGAGCTGAAAGATATGTCAGCTCTTAAAGG 1090
| | | | |
Db 1065 TGGAAATGCTGTTTAAAGGCTCTTCAAAGAGCTGAAAGATATGTCAGCTCTTAAAGG 1124
| | | | |
QY 1091 GTTGGGCTGGATGCCAGGTAAAGTCTTTTGTCTTAAAGAAAGAACTAGAGTCA 1150
| | | | |
Db 1125 GTTGGGCTGGATGCCAGGTAAAGTCTTTTGTCTTAAAGAAAGAACTAGAGTCA 1184
| | | | |
QY 1151 AAATCTGTCCGTGACCTATGATTAATTTTAAAGATGTGCCACTGGCAAAATGTA 1210
| | | | |
Db 1185 AAATCTGTCCGTGACCTATGATTAATTTTAAAGATGTGCCACTGGCAAAATGTA 1244
| | | | |
QY 1211 ACTGTCCAGTCTTTCATATATAAAGGCTTTGAGTAACTCACTGAGGGTATCTGACA 1270
| | | | |
Db 1245 ACTGTCCAGTCTTTCATATATAAAGGCTTTGAGTAACTCACTGAGGGTATCTGACA 1304
| | | | |
QY 1271 ATGCTGAGGTATGAACAAAGTGAAGAAATGAATGATGCTCTTACGAAAAACATG 1330
| | | | |
Db 1305 ATGCTGAGGTATGAACAAAGTGAAGAAATGAATGATGCTCTTACGAAAAACATG 1364
| | | | |
QY 1331 TATGTGATTTCAATCCACGACTTATTAAGAAAGGTGAGTAATTCACAGCTATTTT 1390
| | | | |
Db 1365 TATGTGATTTCAATCCACGACTTATTAAGAAAGGTGAGTAATTCACAGCTATTTT 1424
| | | | |
QY 1391 TGAATATTTTAAATATTTTAAAGATTTCAAGCTATTTCCCTCAATCTGAGGAGC 1450
| | | | |
Db 1425 TGAATATTTTAAATATTTTAAAGATTTCAAGCTATTTCCCTCAATCTGAGGAGC 1484
| | | | |
QY 1451 TGAATATTTTAAATATTTTAAAGATTTCAAGCTATTTCCCTCAATCTGAGGAGC 1504
| | | | |
Db 1485 TGAATATTTTAAATATTTTAAAGATTTCAAGCTATTTCCCTCAATCTGAGGAGC 1544
| | | | |
QY 1505 TTATATGTTGCTATATATAAAGAGTGTCTGC 1536
| | | | |
Db 1545 TTATATGTTGCTATATATAAAGAGTGTCTGC 1576
| | | | |

1	611.8	39.8	18596	US-09-954-531-124	Sequence 124, App
2	611.8	39.8	18596	US-09-954-531-148	Sequence 348, App
3	611.8	39.8	18596	US-09-980-107-1590	Sequence 150, App
4	611.8	39.8	18596	US-09-967-1684-119	Sequence 119, App
5	532	34.6	532	US-10-046-935-1148	Sequence 2148, App
6	340.8	22.2	346	US-09-932-217-7443	Sequence 443, App
7	340.8	22.2	346	US-09-833-263-4443	Sequence 443, App
8	180	9.0	795	US-09-741-669-181	Sequence 181, App
9	137.6	9.0	640681	US-09-750-988-1	Sequence 1, App
10	126.8	8.3	1200	US-09-954-197-1	Sequence 1, App
11	126.8	8.3	2730	US-09-967-446-1	Sequence 1, App
12	116.8	7.7	408	US-09-867-550-395	Sequence 395, App
13	68.2	4.4	213	US-09-974-300-5027	Sequence 5027, App
14	59.2	3.9	4605	US-09-070-9278-315	Sequence 315, App
15	55.4	3.6	57130	US-09-835-081-3	Sequence 3, App
16	52.2	3.4	514	US-09-974-300-5022	Sequence 5022, App
17	52.2	3.4	528	US-09-974-300-5023	Sequence 5023, App
18	51.8	3.4	43058	US-09-954-456-295	Sequence 295, App
19	51.8	3.4	43058	US-09-954-456-529	Sequence 529, App

```

RESULT 1
US-09-954-531-124
; Sequence 124, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cancer
; TITLE OF INVENTION: Gene Sets
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 124
; LENGTH: 18596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-124

```

	Query Match	Best Local Similarity	39.8%;	Score 611.8;	DB 9;	Length 18596;
	Matches 665;	Conservative 0;	Mismatches 65;	Indels 6;	Gaps 1;	
QY	807	GTAACATGATTGGCAGCATCA	CGGCGCTTGAAGCCAGGCGACTTTATACACTTTGGGAGA	866		
Db	15510	GAACCTTGTGTATACATCTCTGTACTTGTTCACGACATAGGAGCATTAACAAAG	15563			
QY	867	TGACATATTTTACTGTATACATCGACCACTGAATAATTCAGCTTCAGGGAACCCAG	926			
Db	15570	GTCGACAAATTATGCAAAATTAATGGCTTATTTTGTTTTATAGCTTCAGGGAACCCAG	15623			
QY	927	ACCTTTCCCAAAGCTCAGATTCTTCGAAAAGTTGAGAAAATGTAGTACTTCAAAGCTGA	986			


```

Query Match      39.8%; Score 611.8; DB 9; Length 18536;
Best Local Similarity 90.4%; Pred. No. 2,5e-154;
Matches 665; Conservative 0; Mismatches 65; Indels 6; Gaps 1.

QY      807  GNACTGATTTGGCCACATCAACGGGCGCTCGAAGCCAGGTGACTTTATATACACACTTTGGAGAGA 866
      |||
Db      15510  GAACCTTGTGTGATCAACATCTCTGTATCTTTGTTCACGACACATGAGAGCAATATCAACAG 15569
QY      867  TGCACATATTTTACCTGAAATCACATGAGCCACCTGAAAAATTCAGCTTCAGCGAACCAG 926
      |||
Db      15570  GTGCGACATTTATAGGCAAAATATAGGCGCTTATTTTGTTTTGTAGCTTCAGCGAACCAG 15629
QY      927  ACCCTTCCCAAGCTGAGATTTCTTCGAAAAGTTGAAAAATTTGAAGATCTCGAAAGCTGA 986
      |||
Db      15630  ACCCTTCCCAAGCTGAGATTTCTTCGAAAAGTTGAAAAATTTGAAGATCTCGAAAGCTGA 15689
QY      987  AAGCTTTCAAGTTGAAAGGACCAATCCGATCCAACTTAAATATGAATAGGCTGTATTA 1046
      |||
Db      15690  AAGCTTTCAAGTTGAAAGGACCAATCCGATCCAACTTAAATATGAATATAGGCTGTATTA 15749
QY      1047  GGGTCCTTTCAAGAGACTGGAAGATATATGCTAGTCTTTAGGGGTTTGGGCTGAAATGCCG 1106
      |||
Db      15750  GGGTCCTTTCAAGAGACTGGAAGATATATGCTAGTCTTTAGGGGTTTGGGCTGAAATGCCG 15809
QY      1107  AGGTAAAGTCTCTTTTGTCTTAAAGANAGAAGACTAGCTCAAAAATCTGTCCGTGAC 1166
      |||
Db      15810  AGGTAAAGTCTCTTTTGTCTTAAAGANAGAAGACTAGCTCAAAAATCTGTCCGTGAC 15869
QY      1167  CTATCAGTTATTAATTTTAAAGATGTGCGCATGCGCAAAATGTATCTGTGCCAGTTCTTT 1226
      |||
Db      15870  CTATCAGTTATTAATTTTAAAGATGTGCGCATGCGCAAAATGTATCTGTGCCAGTTCTTT 15929
QY      1227  CCATTAATTAAGGCTTTGATTACTCACTGAGGGTATCTGACAAATGTATGTGCAATTC 1286
      |||
Db      15930  CCATTAATTAAGGCTTTGATTACTCACTGAGGGTATCTGACAAATGTATGTGCAATTC 15989
QY      1287  CAAAGTGAAGGAAATGAAATGTATGTGCTTAGCAAAAACATGTATGTGCAATTCATTC 1346
      |||
Db      15990  CAAAGTGAAGGAAATGAAATGTATGTGCTTAGCAAAAACATGTATGTGCAATTCATTC 16049
QY      1347  CGACGACTTAAAGAAGGTGTGATTTTCAACAGCTATTTTGGAAATTTTATAGA 1406
      |||
Db      16050  CGACGACTTAAAGAAGGTGTGATTTTCAACAGCTATTTTGGAAATTTTATAGA 16109
QY      1407  TATTTTAAGATTTTCAACAGCTATTCCTCCAAATCTGAGGAGCTGATATACACATGCA 1466
      |||
Db      16110  TATTTTAAGATTTTCAACAGCTATTCCTCCAAATCTGAGGAGCTGATATACACATGCA 16169
QY      1467  TCAATGATGAGAGTGTGTTATGAACT-----TTAATGTTGTTTATATGTTGCTATATA 1520
      |||
Db      16170  TCAATGATGAGAGTGTGTTATGAACTTTAAAGTTATATGTTGTTTATATGTTGCTATATA 16229
QY      1521  TAAAGAAAGTGTCTGC 1536
      |||
Db      16230  TAAAGAAAGTGTCTGC 16245

RESULT 3
US-09-880-107-1590
; Sequence 1590, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Home, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880, 107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054

```


PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1590
LENGTH: 18596
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 D00596
US-09-880-107-1590

Query Match 39.8%; Score 611.8; DB 10; Length 18596;
Best Local Similarity 90.4%; Pred. No. 2.5e-154;
Matches 665; Conservative 0; Mismatches 65; Indels 6; Gaps 1;

```
QY 807 GTACATGATTCGCGACATCAACGCGCCTGAAGCCAGGAGCTTATATACACTTTGGAGCA 866
Db 15510 GAACCTTTGTTGATCAGATCCTGCTGTGATCTGTTTACGAGCAGATGAGACATTAACAAG 15569
QY 867 TGCACATATTTTACCTGATCATCATGACGACCTGAAAATTCAGCTTCGCGAGAACCCAG 926
Db 15570 GTCTGACATTTATGCGAAAATATAGCCTTATTTTGTGTTTATGCTTCAGCGAGAACCCAG 15629
QY 927 ACCCTTCCCAAGCTCAGGATTTCTGAAAAGTTGAGAAAATTTGATGACTTCAAGCTGA 986
Db 15630 ACCCTTCCCAAGCTCAGGATTTCTGAAAAGTTGAGAAAATTTGATGACTTCAAGCTGA 15689
QY 987 AGACTTCAGATTTGAGAGGTGACATCCGACCTCAACTATTAATGAAATGAGCTGTTTA 1046
Db 15690 AGACTTCAGATTTGAGAGGTGACATCCGACCTCAACTATTAATGAAATGAGCTGTTTA 15749
QY 1047 GGGTGTCTTCAAAGAGCTGGAAGATTTGCACTCTTAAAGGGTGTGGCTGAGATGCCG 1106
Db 15750 GGGTGTCTTCAAAGAGCTGGAAGATTTGCACTCTTAAAGGGTGTGGCTGAGATGCCG 15809
QY 1107 AGGTAAAGTCTTTTGTCTTAAAGAAAGAAAGAACTAGGTCMAAAATCTGTCCGTGAC 1166
Db 15810 AGGTAAAGTCTTTTGTCTTAAAGAAAGAAAGAACTAGGTCMAAAATCTGTCCGTGAC 15869
QY 1167 CATACGTTATTAATTTTAAAGATTTGCCACTGCGCAATGTATCTGTGCCAGTTCTTT 1226
Db 15870 CATACGTTATTAATTTTAAAGATTTGCCACTGCGCAATGTATCTGTGCCAGTTCTTT 15929
QY 1227 CCATATTAAGGCTTTGAGTTAACTCACTGAGGTTATCTGACATGCTGAGGTTATGAA 1286
Db 15930 CCATATTAAGGCTTTGAGTTAACTCACTGAGGTTATCTGACATGCTGAGGTTATGAA 15989
QY 1287 CAAAGTGAAGAGATGAATGTATGTCTTTAGCAAAAACATGTATGTGCATTTCAATC 1346
Db 15990 CAAAGTGAAGAGATGAATGTATGTCTTTAGCAAAAACATGTATGTGCATTTCAATC 16049
QY 1347 CCAAGTACTTAAAGAGGTTGAGTTAACTCAAGCTATTTTGGAAATTTTAAAGAA 1406
Db 16050 CCAAGTACTTAAAGAGGTTGAGTTAACTCAAGCTATTTTGGAAATTTTAAAGAA 16109
QY 1407 TATTTTAAAGATTTCAAGGCTATTCCTCAATCTGAGGAGCTGAGTAACCAATCGA 1466
Db 16110 TATTTTAAAGATTTCAAGGCTATTCCTCAATCTGAGGAGCTGAGTAACCAATCGA 16169
QY 1467 TCATGATGTAGAGTGTGTTATGAAGT-----TAAAGTGTGTTTATATGTTGCTATAA 1520
Db 16170 TCATGATGTAGAGTGTGTTATGAAGT-----TAAAGTGTGTTTATATGTTGCTATAA 16229
QY 1521 TAAAGAGTGTCTGC 1536
Db 16230 TAAAGAGTGTCTGC 16245
```

RESULT 4
US-09-967-768A-119
Sequence 119, Application US/09967768A
Patent No. US2002015087A1
GENERAL INFORMATION:

APPLICANT: Augustus, Keena
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
FILE OR INVENTION: Sets
FILE REFERENCE: 689290-72
CURRENT APPLICATION NUMBER: US/09/967,768A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US/60/236,109
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,034
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,111
PRIOR FILING DATE: 2000-09-28
NUMBER OF SEQ ID NOS: 325
SOFTWARE: PatentIn version 3.0
SEQ ID NO 119
LENGTH: 18596
TYPE: DNA
ORGANISM: Homo sapiens
US-09-967-768A-119

Query Match 39.8%; Score 611.8; DB 10; Length 18596;
Best Local Similarity 90.4%; Pred. No. 2.5e-154;
Matches 665; Conservative 0; Mismatches 65; Indels 6; Gaps 1;

```
QY 807 GTACATGATTCGCGACATCAACGCGCCTGAAGCCAGGAGCTTATATACACTTTGGAGCA 866
Db 15510 GAACCTTTGTTGATCAGATCCTGCTGTGATCTGTTTACGAGCAGATGAGACATTAACAAG 15569
QY 867 TGCACATATTTTACCTGATCATCATGACGACCTGAAAATTCAGCTTCGCGAGAACCCAG 926
Db 15570 GTCTGACATTTATGCGAAAATATAGCCTTATTTTGTGTTTATGCTTCAGCGAGAACCCAG 15629
QY 927 ACCCTTCCCAAGCTCAGGATTTCTGAAAAGTTGAGAAAATTTGATGACTTCAAGCTGA 986
Db 15630 ACCCTTCCCAAGCTCAGGATTTCTGAAAAGTTGAGAAAATTTGATGACTTCAAGCTGA 15689
QY 987 AGACTTCAGATTTGAGAGGTGACATCCGACCTCAACTATTAATGAAATGAGCTGTTTA 1046
Db 15690 AGACTTCAGATTTGAGAGGTGACATCCGACCTCAACTATTAATGAAATGAGCTGTTTA 15749
QY 1047 GGGTGTCTTCAAAGAGCTGGAAGATTTGCACTCTTAAAGGGTGTGGCTGAGATGCCG 1106
Db 15750 GGGTGTCTTCAAAGAGCTGGAAGATTTGCACTCTTAAAGGGTGTGGCTGAGATGCCG 15809
QY 1107 AGGTAAAGTCTTTTGTCTTAAAGAAAGAAAGAACTAGGTCMAAAATCTGTCCGTGAC 1166
Db 15810 AGGTAAAGTCTTTTGTCTTAAAGAAAGAAAGAACTAGGTCMAAAATCTGTCCGTGAC 15869
QY 1167 CATACGTTATTAATTTTAAAGATTTGCCACTGCGCAATGTATCTGTGCCAGTTCTTT 1226
Db 15870 CATACGTTATTAATTTTAAAGATTTGCCACTGCGCAATGTATCTGTGCCAGTTCTTT 15929
QY 1227 CCATATTAAGGCTTTGAGTTAACTCACTGAGGTTATCTGACATGCTGAGGTTATGAA 1286
Db 15930 CCATATTAAGGCTTTGAGTTAACTCACTGAGGTTATCTGACATGCTGAGGTTATGAA 15989
QY 1287 CAAAGTGAAGAGATGAATGTATGTCTTTAGCAAAAACATGTATGTGCATTTCAATC 1346
Db 15990 CAAAGTGAAGAGATGAATGTATGTCTTTAGCAAAAACATGTATGTGCATTTCAATC 16049
QY 1347 CCAAGTACTTAAAGAGGTTGAGTTAACTCAAGCTATTTTGGAAATTTTAAAGAA 1406
Db 16050 CCAAGTACTTAAAGAGGTTGAGTTAACTCAAGCTATTTTGGAAATTTTAAAGAA 16109
QY 1407 TATTTTAAAGATTTCAAGGCTATTCCTCAATCTGAGGAGCTGAGTAACCAATCGA 1466
Db 16110 TATTTTAAAGATTTCAAGGCTATTCCTCAATCTGAGGAGCTGAGTAACCAATCGA 16169
QY 1467 TCATGATGTAGAGTGTGTTATGAAGT-----TAAAGTGTGTTTATATGTTGCTATAA 1520
Db 16170 TCATGATGTAGAGTGTGTTATGAAGT-----TAAAGTGTGTTTATATGTTGCTATAA 16229
QY 1521 TAAAGAGTGTCTGC 1536
```

Db 16230 TAAAGAGGTCTTGC 16245

RESULT 5
US-10-046-935-2148

; Sequence 2148, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:

; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.

; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun

; APPLICANT: Stolk, John A.
; APPLICANT: Stolk, John A.

; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER

; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046.935

; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2148

; LENGTH: 532
; TYPE: DNA

; ORGANISM: Homo sapiens
US-10-046-935-2148

Query Match 34.6%; Score 532; DB 9; Length 532;
Best Local Similarity 100.0%; Pred. No. 9.3e-134;
Matches 532; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 203 ACCTGGGGGAGATCCCAACATCCCGCGCGGCGCAGAGAGACGACGAGGCA 262
Db 1 ACCTGGGGGAGATCCCAACATCCCGCGCGGCGCAGAGAGACGAGGCA 60
QY 263 CCGGACCCCTGTGGATTCGGGATTCGGGAGGCGCTACAGAGATGATTCCTC 322
Db 61 CCGGACCCCTGTGGATTCGGGATTCGGGAGGCGCTACAGAGATGATTCCTC 120
QY 323 TGCTGACAAACCAACGCTGTCTTCTGAAAGGCTGTCTTGAAGAGTGTCTGTTATCA 382
Db 121 TGCTGACAAACCAACGCTGTCTTCTGAAAGGCTGTCTTGAAGAGTGTCTGTTATCA 180
QY 383 AGGAGTCCCAAAATGCTAAAGAGCTGTCTTCCAGAGAGTGAAGATCTGGATGCAATG 442
Db 181 AGGAGTCCCAAAATGCTAAAGAGCTGTCTTCCAGAGAGTGAAGATCTGGATGCAATG 240
QY 443 GATCCCGAGCTTTTGGACAGCTTGGATTCCTCACCAGAGAGAGGAGCTTGGGCC 502
Db 241 GATCCCGAGCTTTTGGACAGCTTGGATTCCTCACCAGAGAGAGGAGCTTGGGCC 300
QY 503 CAGTTATGCTTCCAGTGGAGGCAATTTGGGGGAGAAATCAGAGATGATGATCAAT 562
Db 301 CAGTTATGCTTCCAGTGGAGGCAATTTGGGGGAGAAATCAGAGATGATGATCAAT 360
QY 563 ATTGAGGACAGGAGTGAACCAATGCAAGAGATGATGCAACATCAAAACCAACCTG 622
Db 361 ATTGAGGACAGGAGTGAACCAATGCAAGAGATGATGCAACATCAAAACCAACCTG 420
QY 623 AGGACAGAGATGATGATGAGCTTGGATTCAGAGATGATGATGATGATGATGATGAT 682
Db 421 AGGACAGAGATGATGATGAGCTTGGATTCAGAGATGATGATGATGATGATGATGAT 480
QY 683 CTCATGCGATGCGCTTCTGCGAGTGTCTATGTGTGAACAGTGAAGCTGTCTCTG 734
Db 481 CTCATGCGATGCGCTTCTGCGAGTGTCTATGTGTGAACAGTGAAGCTGTCTCTG 532

RESULT 6
US-09-922-217-443/C
; Sequence 443, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Yugu
; APPLICANT: Stolk, John A.
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon B.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922.217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 443
; LENGTH: 346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-217-443

Query Match 22.2%; Score 340.8; DB 10; Length 346;
Best Local Similarity 98.8%; Pred. No. 2.4e-82;
Matches 342; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1007 ACATCCGATCCACTATTAATGGAATGCTGTTAGGAGTGTCTTCAAGAGCTN 1066
Db 346 ACATCCGATCCACTATTAATGGAATGCTGTTAGGAGTGTCTTCAAGAGCTN 287
QY 1067 GAAGATATGCTAGCTTTAGGAGTGTGAGTCCAGATGCAAGATGATTTTCT 1126
Db 286 GAAGATATGCTAGCTTTAGGAGTGTGAGTCCAGATGCAAGATGATTTTCT 227
QY 1127 CTAAAG 1186
Db 226 CTAAAG 167
QY 1187 AGGATGTTCCACTGCGCAATGTAATGCTGTCAGATTCCTTCCATTAATTAAGCTTTGAG 1246
Db 166 AGGATGTTCCACTGCGCAATGTAATGCTGTCAGATTCCTTCCATTAATTAAGCTTTGAG 107
QY 1247 TTACTGCTAGAGGTATCTGACATGCTGAGCTTATGAGAGAGAGAGAGAGAT 1306
Db 106 TTACTGCTAGAGGTATCTGACATGCTGAGCTTATGAGAGAGAGAGAGATGAGAT 47
QY 1307 GATATGCTTTAGCAAAACATGATGATGATGATGATGATGATGATGATGATGAT 1352
Db 46 GATATGCTTTAGCAAAACATGATGATGATGATGATGATGATGATGATGATGAT 1

RESULT 7
US-09-833-263-443/C
; Sequence 443, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:

; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833.263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 443
; LENGTH: 346
; TYPE: DNA

```
! ORGANISM: Homo sapiens
US-09-833-263-443

Query Match
Best Local Similarity 98.8%; Score 340.8; DB 10; Length 346;
Pred. No. 2.4e-82;
Matches 342; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1007 ACAATCCGATCCACTATTAATGAAATGGCTGTTAGGCTCTTCAAGAGCTN 1066
DB 346 ACAATCCGATCCACTATTAATGAAATGGCTGTTAGGCTCTTCAAGAGCTT 287
QY 1067 GAAGGATATGTGCTTAAAGGGTGGGCTGGATGCCAGAGTAAAGTCTTTTGT 1126
DB 286 GAAGGATATGTGCTTAAAGGGTGGGCTGGATGCCAGAGTAAAGTCTTTTGT 227
QY 1127 CTAAGAGANAAGAACTAGTCAAAAATCTGTCCTGACCTATCAATTAATTTTA 1186
DB 226 CTAAGAGANAAGAACTAGTCAAAAATCTGTCCTGACCTATCAATTAATTTTA 167
QY 1187 AGATGTTGCCACTGSCAATGTAACTGTGCCAGTCTTCCATTAATAAAGCTTGGG 1246
DB 166 AGATGTTGCCACTGSCAATGTAACTGTGCCAGTCTTCCATTAATAAAGCTTGGG 107
QY 1247 TTAATCACTGAGGATCTGATCTGATGCTGATGATGATGATGATGATGATGAT 1306
DB 106 TTAATCACTGAGGATCTGATCTGATGCTGATGATGATGATGATGATGATGAT 47
QY 1307 GATGCTCTTACGAAAACATGATGATGATGATGATGATGATGATGATGATGAT 1352
DB 46 GATGCTCTTACGAAAACATGATGATGATGATGATGATGATGATGATGATGAT 1

RESULT 8
US-09-741-669-181
! Sequence 181, Application US/09741669
! Patent No. US20020227181A1
! GENERAL INFORMATION:
! APPLICANT: Forsyth, R. Allen
! APPLICANT: Ohlsen, Karl L.
! APPLICANT: Zyskind, Judith W.
! TITLE OF INVENTION: Genes identified as required for
! proliferation of E. coli
! FILE REFERENCE: ELITRA.009A
! CURRENT APPLICATION NUMBER: US/09/741,669
! PRIOR FILING DATE: 2000-12-19
! PRIOR APPLICATION NUMBER: US 60/173005
! PRIOR FILING DATE: 1999-12-23
! NUMBER OF SEQ ID NOS: 481
! SOFTWARE: FastSeq for Windows Version 4.0
! SEQ ID NO 181
! LENGTH: 795
! TYPE: DNA
! ORGANISM: Escherichia coli
! NAME/KEY: CDS
! LOCATION: (1)...(795)
US-09-741-669-181

Query Match
Best Local Similarity 54.0%; Score 180; DB 10; Length 795;
Pred. No. 8.6e-39;
Matches 458; Conservative 0; Mismatches 330; Indels 60; Gaps 2;

QY 189 CAGTACCTGGGCGATCCACACATCTCCGCTGGCGGCTCAAGAGAGAGAGAGAG 258
DB 7 CAGTATTTTGAATGATGCAAAAAGTCTGCAAGAGAGAGAGAGAGAGAGAGAGAG 66
QY 259 GGCACCGGACCTGTCGATTTGGCATGACAGCGCGCTTACAGGCTGAGATGAATTC 318
DB 67 GGCACCGGACCTGTCGATTTGGCATGACAGCGCGCTTACAGGCTGAGATGAATTC 126
QY 319 CCTGCTGACACACAAAGTGTGCTGGAAGGAGTGTGGAAGAGAGTGTGAGTGT 378
DB 127 CCTGCTGACACACAAAGTGTGCTGGAAGGAGTGTGGAAGAGAGTGTGAGTGT 186
```

```
QY 379 ATCAAGGATCCACAAATGCTAAAGAGCTGTCTTCCAGAGGAGTGAATTCGGAGTCC 438
DB 187 CTCAGAGGAGGACACTTAACATCTGATCTTACACAGAAAACANTGTCACTCTGGAGAG 246
QY 439 AATGATCCGAGACCTTTTGGACAGCGCTGGAGATTTCTCCACAGAGAAAGAGGACT 498
DB 247 TGGG-----CCAGTAAACCGGACCTC 270
QY 499 GGCCAGATTATGCTTCCAGTGAAGAGATTTGGGCGAGATAACAGAGATATGGAATCA 558
DB 271 GGCCAGATTATGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 315
QY 559 GATTAATCAGACAGGAGTGAACCACTGCAAAAGTGAATGACACATCAAAACCAAC 618
DB 316 -----GGTCTCAATTAATGACAGATCACTAGGTATGAACAGGTGAAGAGAGAG 356
QY 619 CCTGACAGACAAATATCATGATGCTGCTGATGATGATGATGATGATGATGATGATG 678
DB 367 CCTGACAGACAAATATCATGATGCTGCTGATGATGATGATGATGATGATGATGATG 426
QY 679 CTGCTTCATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 728
DB 427 CTGCTTCATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 486
QY 739 CTGACAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 798
DB 487 CTGACAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
QY 799 CTGCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 858
DB 547 CTGCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 606
QY 859 TTGGAGATGACAAATATTAATGATGATGATGATGATGATGATGATGATGATGAT 918
DB 607 TTGGAGATGACAAATATTAATGATGATGATGATGATGATGATGATGATGATGAT 666
QY 919 GAACCAAGCTTTCCAAAGCTCAGATTTCTGAAAAGTGAAGAAATGATGATGAT 978
DB 667 GAACCAAGCTTTCCAAAGCTCAGATTTCTGAAAAGTGAAGAAATGATGATGAT 726
QY 979 AAGGTGAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1038
DB 727 AAGGTGAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
QY 1039 GCTGTTA 1046
DB 787 GCTGTTA 794

RESULT 9
US-09-790-988-1
! Sequence 1, Application US/09790988
! Patent No. US20020127687A1
! GENERAL INFORMATION:
! APPLICANT: SHIGEMOTO, SHUJI
! APPLICANT: WATANABE, HIDEMI
! APPLICANT: HATORI, MASAHIRA
! APPLICANT: SAKAKI, YOSHIYUKI
! TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
! FILE REFERENCE: 081356/0159
! CURRENT APPLICATION NUMBER: US/09/790,988
! PRIOR FILING DATE: 2001-02-23
! PRIOR APPLICATION NUMBER: JP2000-107160
! PRIOR FILING DATE: 2000-04-07
! NUMBER OF SEQ ID NOS: 7
! SOFTWARE: PatentIn Ver. 2.1
! SEQ ID NO 1
! LENGTH: 640681
! TYPE: DNA
! ORGANISM: Buchnera sp.
US-09-790-988-1
```

Query Match 9.0%; Score 137.6; DB 10; Length 640681;
Best Local Similarity 55.4%; Pred. No. 1.3e-25;
Matches 266; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

```
QY 567 AGGACAGGAGTTGACCACTGCAAGAGTGTGACACCATCAAAAACCAACCTGACGA 626
Db 480465 AGGACAGCAATTCATCAATTAATAATGATTTTAATCAATTAATAATAATCCGATTC 480524
QY 627 CAGAAATCATCATGTGCGCTTGGAATCCAGAGATCTTCTCTGATGCGCGCTGCTCC 666
Db 480525 CCGTAGAATGTAGTTTCTAGTTGGAATGTTGAGATATAGATAAATGAGATTACCTCC 480584
QY 687 ATGCGATGCCCTCTGCGAGTTCTATGTGTGTAACAATGAGCTGTCTCTGCGAGTACCA 746
Db 480585 TTGCGATGTTCTATTTCAATTTATATGATTAACAATAATTAAGTTGCAATGTACCA 480644
QY 747 GAGATCGGAGACATGGGCGCTCGGTGTGCTTTCAACATGCGCAGCTAGCGCTGCTCAC 806
Db 480645 GCGTCTTGTGATGTGTTTCTAGGACTACCTTTTAATATAGTAACTCAATTAATTAT 480704
QY 807 GTACATGATTCGACATCAACGCGGCTGAAGCAGTGACTTATACACACTTTGGGAGA 866
Db 480705 AACATGATAGCAACAACATGTGATTTAAAGTTGAGATTTTGTGACAGAGGCGCA 480764
QY 867 TGCAATATTTACCTGATCATCATGAGCGCACTGAAAATTCAGCTTCAGCGAGAACCCAG 926
Db 480765 TGTTCATCTTATATATATATATATATATATATATATATATATATATATATATATAT 480824
QY 927 ACCCTTCCAAAGCTCAGATTTCTTGAAAAGTTGAGAAAATTTGATGACTTCAAAAGCTGA 986
Db 480825 AACGTTCCCAACTAATCAATCTCTCAAAAACCTCAGTCAATATTTCAATATTTTTCGA 480884
QY 987 AGACTTTCAATTTGAAGGTTACATTCGCAATCCCACTATTAATAATGGAATGGCTGTTTA 1046
Db 480885 AGACTTTAAATTTATGATATCATCTTATCTGCTATCAAGAGGAATATCTAATATA 480944
```

RESULT 10

```
US-09-954-197-1
; Sequence 1, Application US/09954197
; Patent No. US20020107379A1
; GENERAL INFORMATION:
; APPLICANT: MARK, Achim
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR THE thyA GENE
; FILE REFERENCE: 032301 WD 201
; CURRENT APPLICATION NUMBER: US/09/954,197
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (231)..(1028)
; OTHER INFORMATION:
US-09-954-197-1
```

Query Match 8.3%; Score 126.8; DB 10; Length 1200;
Best Local Similarity 56.7%; Pred. No. 2.4e-24;
Matches 254; Conservative 0; Mismatches 191; Indels 3; Gaps 1;

```
QY 578 TTGACCACTGCAAAAGTGAATTGACACCATCAAAACCACTGACGACAGAAATCA 637
Db 565 TTGACCAATCTCAGGTGCTTTAGAACTCTGCGAAACACCTGATTCACCTGCAATA 624
QY 638 TCAATGCGCTTGGAATCCAAAGATCTTCCCTGATGCGCGCTCCATCCATGCGCC 697
Db 625 TTGCTCGGCGGTGGAATGTTTCCGAGCTTGAAGAAACATGAGCTTTCCCTTGTCACTTGC 684
QY 698 TCTGCAAGTCTATGTGTGAACAAGTGAAGCTGTCTCTGCAAGCTGTAACAGAGATGGAG 757
```

```
Db 685 TTTTCAGACTCTATGTGCGCAGATGGCAACTGTTTGCAGCTTACCAAGCTTCTGCG 744
QY 758 ACATGGGCGCTGCGTGTGCGCTTTCAACATGCGCAGCTAGCGCTGCTCACTGATGTTG 817
Db 745 ACATGTTCTGCGGTGTGCGCTTTCAACATGCGCAGCTTATGCACTGCTTCCACCAATGTTG 804
QY 818 CGCATCAAGGCGCTGGAAGCGAGTGAATTTATACACTTTTGGAGATGACATATTT 877
Db 805 CCGACGAGGAGGCTTGAAGTGTGCGAGATTTGATTTGAGCTGCGGCGAGCCACATTT 864
QY 878 ACTGAAATCATGAGCCACTGAAAATTCAGCTTCAAGGAGAAACCCAGACTTTCCCA 937
Db 865 ATGACAAACCAAGAAAGAGTGTGCGGAGCAGCTGAGCCGCGAAGCTGCCCCAATCCCA 924
QY 938 AGCTGAGATTTTGAAAAGTTGAAAATTAATGATGACTTCAAACTGAATGACTTTGAGA 997
Db 925 ---CCTTGAGCTCAACAAGGCGAGCTCCATGTTTGAATGATGAGCTTGATGATCAACCG 981
QY 998 TTGAAGGTACATCGGCTCCACTAT 1025
Db 982 TGTCCGCGTACGATCCACACCATTTGAT 1009
```

RESULT 11

```
US-09-987-446-1
; Sequence 1, Application US/09987446
; Patent No. US2002011562A1
; GENERAL INFORMATION:
; APPLICANT: FARWICK, Mike, et al.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR THE cysQ GENE
; FILE REFERENCE: 032301 WD 245
; CURRENT APPLICATION NUMBER: US/09/987,446
; CURRENT FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2730
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1014)..(1769)
; OTHER INFORMATION:
US-09-987-446-1
```

Query Match 8.3%; Score 126.8; DB 10; Length 2730;
Best Local Similarity 52.0%; Pred. No. 3.9e-24;
Matches 407; Conservative 0; Mismatches 312; Indels 63; Gaps 3;

```
QY 244 AAGACGACCGGACGCGACCGGACCGCTGCGATTTGGCATGAGCGCGCTACAGC 303
Db 1851 AAGACGACCGGACGCGACCGGACCGGACCTACTTTATTCGACAAATTCGCTTTGAT 1910
QY 304 CTGAGAGATGATCCCTCTGCTGACAAACCAAGCTGTCTTGAAGAGTGTGTTGAG 363
Db 1911 CTGATGAAAGTGTTCCTCTTCTGACCAAGAAAGTTCATTTCACTCTGTGTGGGT 1970
QY 364 GAGTGTGCTGTTTATCAGGATTCACAAATGCTTAAAGCTGTCTTCCAAAGGAGTG 423
Db 1971 GAGCTTTGTGTGCTTCTTCAAGGAGATTCACAGTCAAAATGCTGAGATTAACAATC 2030
QY 424 AAATCTGGAGTGCATATGATCCGAGATCTTTTGAACAGCTTGGATTTCCACAGCA 483
Db 2031 CGATTTGGAATGATGAGGAGAT----- 2054
QY 484 GAAAGAGGAGCTTGGGCGCAATTTATGCTTCCAGTGAAGGATTTTGGGCGAGATAC 543
Db 2055 GAGGACGCGGAGCTGGGCGCTGTTTATGTGTCCAGTGGGCTTCTTGGC----- 2103
QY 544 AAGATATGAAATCAATTAATTCAGGACAGGAGTGAACCACTGAAGAAAGATGATGAC 603
Db 2104 -----CAACCTGATGTGTCTACATTTGACAGATCTCAGGTGCTTTAGAA 2150
```

QY 604 ACCATCAAAACCAACCTGAGACAGAGAAATCATCATGTGCGCTTGAATCCAGAGAT 663
DB 2151 ACTCTGCGAAACACCTGATTCACGTCGCAATATGCTCGGCGTGAATGTTCCGAG 2210
QY 664 CTTCCTGTGATGCGCGCTGCTCCCTCCCTGTCGCACTTCCAGATTCATGTGTGACAGT 723
DB 2211 CTTGAAACATGAGCTCTCCCTCCCTGTCGCACTTCCAGATTCATGTGTGACAGT 2270
QY 724 GAGCTGTCTGCGCACTGTCACAGAGATGCGGAGACATGAGCTCCGCTGCTGCTTCAAC 783
DB 2271 AAATGCTTGGCACTGTCACAGAGATGCGGAGACATGAGCTCCGCTGCTGCTTCAAC 2330
QY 784 ATGCGCACTGAGCGCTGCTGTCACAGATGTCGCGACATGAGCGCTGCGAGAGT 843
DB 2331 ATGCACTTGTATGACAGCTGTCACAGATGTCGCGAGAGATGAGCTGCGAGAGT 2390
QY 844 GACTTATACACACTTGGGAGATGACACATATTTACCTGATACATGAGCGACCTGAAA 903
DB 2391 GAGTTCATTTGAGCTGCGGCGGAGCTGCGACATTTTATGACACACAGAGAGAGT 2450
QY 904 ATTGAGCTTCAGCGAGAACCCAGACCTTTCCCAAGCTCAGAGATTCCTGAAAAGTTGAG 963
DB 2451 GAGCACTGAGCGCGGAGAGCTGCGCTTACCCCA---CCTTGAAGCTCAAGAGAGAG 2507
QY 964 AAATGATGACTTCAAAAGCTGAAAGCTTCAGATTTAGAGGTTACAAATCCGATCCAACT 1023
DB 2508 TCCATGTTTGTAGTACAGCTTCGATGACATCACCGCTGCTCGGCTACGATCCACCACTATG 2567
QY 1024 AT 1025
DB 2568 AT 2569

RESULT 12

US-09-867-550-395
; Sequence 395, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Foad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 395
; LENGTH: 408
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(2)
; OTHER INFORMATION: wherein any n is one of a or t or g or c
US-09-867-550-395

Query Match 7.7%; Score 118.8; DB 10; Length 408;
Best Local Similarity 56.3%; Pred. No. 1.8e-22;
Matches 222; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

QY 524 GGCATTTTGGGCGAGATACAGATGATGAAATGATTTATTCAGACAGGAGTTGACC 583
DB 14 GGGCTAATTTGGCGCACTAAATGCCATGTGACTTAAACCAAGATGTTTGACC 73
QY 584 AACTGCAAGAGTGAATGACACCATCAAAACCACTGACAGACAGAAAGATCATGAT 643
DB 74 AAATTAATGTGTTAATTAATGAATTAATAACCAACCAATTCAGCGCTTAAATGCT 133

QY 644 GCGCTTGGATTCAGAGATCTTCCCTGATGAGCGCTGCTTCATGACAGTCCCTGCGC 703
DB 134 CTGTTGGATTCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 193
QY 704 AGTTCTATGTGTGAAACAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 763
DB 194 AATTTTGTGTAAGAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 253
QY 764 GCGCTGAGTGGCTTCAACATGCGACAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCT 823
DB 254 TCTTATGAGGCTGCTTAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 313
QY 824 TACGCGGCTGAG 883
DB 314 TGTGTGCTTGT 373
QY 884 ATCAGATGAG 917
DB 374 ACCATTTGAAACAGGCAAACTACATTTGACGCG 407

RESULT 13

US-09-974-300-5027
; Sequence 5027, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods for Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5027
; LENGTH: 213
; TYPE: DNA
; ORGANISM: Bacillus clausii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(213)
; OTHER INFORMATION: n = A,T,C or G
US-09-974-300-5027

Query Match 4.4%; Score 68.2; DB 10; Length 213;
Best Local Similarity 57.3%; Pred. No. 5.2e-09;
Matches 121; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 833 TGAAGCAGGTGACTTTATACAGCTTTGGAGATGACATATTTACTGATACATGCG 892
DB 3 TAAAGCAGGTGACTTTATACAGCTTTGGAGATGACATATTTACTGATACATGCG 62
QY 893 AGCAGTGAATTTGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 952
DB 63 AACAGACAGAGCTTCAATTTAGCGCTAAACCAAAACAGCTGCAACAGATGATTCATC 122
QY 953 GAAAGTTGAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1012
DB 123 CAGAGTTAAGAGAGCTGTTTGGTTTACTTACAGATGATTTTGAACAGAAATATGATC 182
QY 1013 CGATTCAGATTTAAATGAGAAATGAGTGT 1043
DB 183 CTCATCCGATATTTAAAGAGAGAGTCCGT 213

RESULT 14

US-09-070-927A-315/c

Sequence 315, Application US/09070927A
Patent No. US20020120116A1
GENERAL INFORMATION:
APPLICANT: Charles A. Kunsch
Patrick J. Dillon
Steven Barash
TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 982
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,927A
FILING DATE: 04-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 315:
SEQUENCE CHARACTERISTICS:
LENGTH: 4605 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 315:
US-09-070-927A-315
Query Match 3.9%; Score 59.2; DB 10; Length 4605;
Best Local Similarity 64.7%; Pred. No. 6.9e-06;
Matches 88; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
QY 578 TTGACCACTGCAAGAGTATGACACATCAAAACCAACCCGACGAGAGAAATCA 637
DB 157 TCGATCAATTAAGCTAATGATGATTAATAAACCACATCTGATCTCGCCGTTAA 98
QY 638 TCATGTGGCTTGGATCGAAGATCTTCTGATGACGCTGCTCCATGCCAGGCC 697
DB 97 TTGTTTCGCTTGAATCCAGAAATGCTTCAATGCGCTTGCCGCTGTGCACACTA 38
QY 698 TCTGCCAGTTCTATGT 713
DB 37 TGTTCATTTATGT 22
RESULT 15
US-09-835-081-3
Sequence 3, Application US/09835081
Patent No. US20020151020A1
GENERAL INFORMATION:
APPLICANT: YAN, Xianghe et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
THEROF

FILE REFERENCE: C1001224
CURRENT APPLICATION NUMBER: US/09/835,081
CURRENT FILING DATE: 2001-04-16
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 57130
TYPE: DNA
ORGANISM: Human
US-09-835-081-3
Query Match 3.6%; Score 55.4; DB 10; Length 57130;
Best Local Similarity 54.7%; Pred. No. 0.00043;
Matches 110; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 1 GGGGGGGGGGGGACCACTTGCTGCTCCGTCCTCCGCGGCGCACTTGGCTGCTCCGT 60
DB 2624 GGGGGGGGGGGGCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2683
QY 61 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
DB 2684 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2743
QY 121 TCGAGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
DB 2744 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2803
QY 181 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 201
DB 2804 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2824

Search completed: November 29, 2002, 00:37:39
Job time : 972.818 secs


```

QY 61 CCGCGCGCCACTTGGCTGCTCCGTCCCGCCCGCCGCGCATGCTGTGGCCGCG 120
DB 61 CCGCGCGCCACTTGGCTGCTCCGTCCCGCCCGCCGCGCATGCTGTGGCCGCG 120
QY 121 TCGGAGCTGCGCGCGCGCGCTTGGCCCGCCCGCGCAAGAGGCGGAGCGCGCGGT 180
DB 121 TCGGAGCTGCGCGCGCGCGCTTGGCCCGCCCGCGCAAGAGGCGGAGCGCGGT 180
QY 181 CCGCGCGCGAGGAGCTGACAGTACCTGAGGAGATCCAGACATCCCTGCGCGCGGT 240
DB 181 CCGCGCGCGAGGAGCTGACAGTACCTGAGGAGATCCAGACATCCCTGCGCGGT 240
QY 241 AGGAGAGAGAGCCGACCGGAGCCGCGACCTGTGCTATTCGCGATGACGCGCGCTAC 300
DB 241 AGGAGAGAGAGCCGACCGGAGCCGCGACCTGTGCTATTCGCGATGACGCGCGCTAC 300
QY 301 AGCCTGAGAGATTTCCCTCTGCTGACCAACCAACGTGTGTTCTGGAAGGCTGTG 360
DB 301 AGCCTGAGAGATTTCCCTCTGCTGACCAACCAACGTGTGTTCTGGAAGGCTGTG 360
QY 361 GAGGAGTGTGCTGTGTTATTCAGAGGATCCAAATGCTAAAGAGCTGTCTTCAAGGGA 420
DB 361 GAGGAGTGTGCTGTGTTATTCAGAGGATCCAAATGCTAAAGAGCTGTCTTCAAGGGA 420
QY 421 GTGAAATCTGAGGATGCGAATGATCCCGAGACTTTTGGACAGCTGCGGATTCGCAAC 480
DB 421 GTGAAATCTGAGGATGCGAATGATCCCGAGACTTTTGGACAGCTGCGGATTCGCAAC 480
QY 481 AGGAGAGAGGAGGAGCTTGGGCGCGAGTTATGAGCTTCAAGTGAAGGATTTGGGCGAGAA 540
DB 481 AGGAGAGAGGAGGAGCTTGGGCGCGAGTTATGAGCTTCAAGTGAAGGATTTGGGCGAGAA 540
QY 541 TACAGAGATGATGATGATGATTTTACAGACAGGAGTTTGAACAACTGCAAGAGTGAAT 600
DB 541 TACAGAGATGATGATGATGATTTTACAGACAGGAGTTTGAACAACTGCAAGAGTGAAT 600
QY 601 GAGACATCAAAACCAACCTGACGACAGAGAGATCATCATGTCGCTTGAATCCAGAA 660
DB 601 GAGACATCAAAACCAACCTGACGACAGAGAGATCATCATGTCGCTTGAATCCAGAA 660
QY 661 GATCTTCTCTGATGAGGCGCTGCTCCGATGCGCATGCGCTTGGCTTCTATGAGTGAAC 720
DB 661 GATCTTCTCTGATGAGGCGCTGCTCCGATGCGCATGCGCTTGGCTTCTATGAGTGAAC 720
QY 721 AGTGAAGTCTGCTGACGCTGCTACAGAGATGCGGAGCATGCGGCTCGGTGCTTTC 780
DB 721 AGTGAAGTCTGCTGACGCTGCTACAGAGATGCGGAGCATGCGGCTCGGTGCTTTC 780
QY 781 AACATGCGCAGCTAGCGCTGCTGCTACAGTATGATGCGCATCAAGGCGCTGAAGCCA 840
DB 781 AACATGCGCAGCTAGCGCTGCTGCTACAGTATGATGCGCATCAAGGCGCTGAAGCCA 840
QY 841 GGTGACTTTATACACATTTTGGAGATGACATATTTACCTGAATCAATCGACCACTG 900
DB 841 GGTGACTTTATACACATTTTGGAGATGACATATTTACCTGAATCAATCGACCACTG 900
QY 901 AAAATTCAGCTTCAAGGAGACCCAGACCTTCCAAAGCTCAGAGATCTTGGAAAAGTT 960
DB 901 AAAATTCAGCTTCAAGGAGACCCAGACCTTCCAAAGCTCAGAGATCTTGGAAAAGTT 960
QY 961 GAGAAATGATGATCTTCAAGGAGCTTCAAGTTCAGATGGAAGGATCAATCCGATCCA 1020
DB 961 GAGAAATGATGATCTTCAAGGAGCTTCAAGTTCAGATGGAAGGATCAATCCGATCCA 1020
QY 1021 ACTATTAATGAAATGCTGTTAGGCTGCTTCAAGGAGCTTGAAGGATATTTGCA 1080
DB 1021 ACTATTAATGAAATGCTGTTAGGCTGCTTCAAGGAGCTTGAAGGATATTTGCA 1080
QY 1081 GTCTTAAGGAGTGGCTGAGATGCGAGGATTAAGTCTTTTGTCTGTAAGAGANAGAG 1140
DB 1081 GTCTTAAGGAGTGGCTGAGATGCGAGGATTAAGTCTTTTGTCTGTAAGAGANAGAG 1140
QY 1141 AACTAGTCAAAATCTGTGCTGATCCTATCAGTTATTAATTTTAAGATGTGCACT 1200

```

```

DB 1141 AACTAGTCAAAATCTGTGCTGATCCTATCAGTTATTAATTTTAAGATGTGCACT 1200
QY 1201 GGCATATGAACTGTGCAGTCTTTCATTAATAAAGGCTTGAATTAATCACTGAG 1260
DB 1201 GGCATATGAACTGTGCAGTCTTTCATTAATAAAGGCTTGAATTAATCACTGAG 1260
QY 1261 GTATCTGACAAATGCTGAGTTATGAACAAAGTGAAGGAGATGAATGATGCTTGA 1320
DB 1261 GTATCTGACAAATGCTGAGTTATGAACAAAGTGAAGGAGATGAATGATGCTTGA 1320
QY 1321 CAAAAACATGATGATGATTCATTCACAGTATTAATAAAGGATGATGATTTGAC 1380
DB 1321 CAAAAACATGATGATGATTCATTCACAGTATTAATAAAGGATGATGATTTGAC 1380
QY 1381 AAGCTATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 1440
DB 1381 AAGCTATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 1440
QY 1441 CTGAGGAGCTGAGTAAACCATCATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1441 CTGAGGAGCTGAGTAAACCATCATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1501 TGTATTATATGTTGCTATTAATTAAGAGTGTTCG 1536
DB 1501 TGTATTATATGTTGCTATTAATTAAGAGTGTTCG 1536

RESULT 3
US-09-347-878-29
; Sequence 29, Application US/09347878C
; Patent No. 6376210
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
; FILE REFERENCE: 25885-1651
; CURRENT APPLICATION NUMBER: US/09/347,878C
; CURRENT FILING DATE: 1999-07-06
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 29
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Human thymidylate synthase gene: exons 1-8
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: D00596/GenBank
; US-09-347-878-29

Query Match 61.3%; Score 942; DB 4; Length 942;
Best Local Similarity 100.0%; Pred. No. 2.9e-241;
Matches 942; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 106 ATGCTGTGCGCGCTGCGAGCTGCGCGCGCGCGCTTGGCCCGCGCAGAGACGG 165
DB 1 ATGCTGTGCGCGCTGCGAGCTGCGCGCGCGCGCGCTTGGCCCGCGCAGAGACGG 60
QY 166 GAGCGCGAGCGCGCTGCGCGCGCGCGCGAGTGTGATGATCTGCGGCGCATATC 225
DB 61 GAGCGCGAGCGCGCTGCGCGCGCGCGCGAGTGTGATGATCTGCGGCGCATATC 120
QY 226 CTCGCGCGCGCGCTGCGAGAGCGACCGGAGCGGAGCGGAGCGGAGTGTGATGAT 285
DB 121 CTCGCGCGCGCGCTGCGAGAGCGACCGGAGCGGAGCGGAGCGGAGTGTGATGAT 180
QY 286 ATGCAAGGCGCTACAGCTGAGAGATGAATTCCTGCTGACACCAACGTTGTC 345
DB 181 ATGCAAGGCGCGCTACAGCTGAGAGATGAATTCCTGCTGACACCAACGTTGTC 240
QY 346 TGAAGAGGTGTTTGAAGAGTGTGCTGTGTTATCAAGGATCCAAATGCTAAAGAG 405
DB 241 TGAAGAGGTGTTTGAAGAGTGTGCTGTGTTATCAAGGATCCAAATGCTAAAGAG 300

```

QY 406 CTGCTTCCAGAGGAGTGAATCTGGATGCAATGATCCGAGACTTTTGGACAGC 465
DB 301 CTGCTTCCAGAGGAGTGAATCTGGATGCAATGATCCGAGACTTTTGGACAGC 360
QY 466 CTGGAGTCTCCACCAAGAGAGAGGAGCTTGGCCCACTTTATGCTTCAATGAGG 525
DB 361 CTGGAGTCTCCACCAAGAGAGAGGAGCTTGGCCCACTTTATGCTTCAATGAGG 420
QY 526 CATTTTGGGCGAGATACAGAGATATGATCATATTTTACAGACAGGAGTTGACCA 585
DB 421 CATTTTGGGCGAGATACAGAGATATGATCATATTTTACAGACAGGAGTTGACCA 480
QY 586 CTGCAAGAGATTTGACACCATCAAAACCACTTACAGACAGAGATATGATGAC 645
DB 481 CTGCAAGAGATTTGACACCATCAAAACCACTTACAGACAGAGATATGATGAC 540
QY 646 GCTTGGATCCAGAGATCTTCTGATGAGGCTGCTCCATGCTCATGCTTGGCAG 705
DB 541 GCTTGGATCCAGAGATCTTCTGATGAGGCTGCTCCATGCTCATGCTTGGCAG 600
QY 706 TTCTATGATGAGACAGTGAATGCTGCTCCAGCTGTAACAGAGATGAGAGATGAGC 765
DB 601 TTCTATGATGAGACAGTGAATGCTGCTCCAGCTGTAACAGAGATGAGAGATGAGC 660
QY 766 CTGCTGATGCTTTCAACATGCGACAGTACGCTGCTCACTGATCATATTTGCGACATC 825
DB 661 CTGCTGATGCTTTCAACATGCGACAGTACGCTGCTCACTGATCATATTTGCGACATC 720
QY 826 ACGGCTCTGAGACAGTGAATTTATACACATTTTGGAGATGACATATTTTCTGAT 885
DB 721 ACGGCTCTGAGACAGTGAATTTATACACATTTTGGAGATGACATATTTTCTGAT 780
QY 886 CACATGAGCCACTGAAATTTAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 945
DB 781 CACATGAGCCACTGAAATTTAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 946 ATTTCTGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1005
DB 841 ATTTCTGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 1006 TACATGCGCATCCACTATTTAAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1047
DB 901 TACATGCGCATCCACTATTTAAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 942

RESULT 4

US-08-965-048-5/c
; Sequence 5, Application US/08965048
; Patent No. 6323244
; GENERAL INFORMATION:
; APPLICANT: Chen, Hong
; APPLICANT: Freimer, Nelson
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF NEUROPSYCHIATRIC DISORDERS
; FILE REFERENCE: 7853-093
; CURRENT APPLICATION NUMBER: US/08/965,048
; CURRENT FILING DATE: 1997-11-05
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 45716
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-965-048-5

Query Match 40.9%; Score 627.8; DB 4; Length 45716;
Best Local Similarity 91.1%; Pred. No. 9,4e-157;
Matches 665; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
QY 807 GTACATGATTCGACATCAAGGCTGAGAGCGAGTGAATTTATACACATTTGGAGAG 866
DB 44683 GAACCTTTGATGATCAATCCCTGCTGATCTTTTCAATGACATGAGAGCAATTACAACAG 44624

QY 867 TGAACATTTTACTGATATCATGAGCCACTGAAATTCAGCTTCAAGAGAGACCCAG 926
DB 44623 GTGATCAATATTATGCGAAATATGAGCTTATTTTGTCTTACCTTCAAGAGACCCAG 44564
QY 927 ACCCTTCCCAAGAGTGAATTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 986
DB 44563 ACCCTTCCCAAGAGTGAATTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 44504
QY 987 AGATTTTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1046
DB 44503 AGATTTTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 44444
QY 1047 GGGTCTTCAAGAGAGTGAATTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1106
DB 44443 GGGTCTTCAAGAGAGTGAATTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 44384
QY 1107 AGGTAAAGATTTTCTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1166
DB 44383 AGGTAAAGATTTTCTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 44324
QY 1167 CTATCAGTATTTATTTTAAAGATGCTGCACTGCAATGATGATGCTGCAATG 1226
DB 44323 CTATCAGTATTTATTTTAAAGATGCTGCACTGCAATGATGATGCTGCAATG 44264
QY 1227 CCAATTTAAAGAGTGAATTTCACTGAGAGGATCTGCAATGCTGAGAGATTTAGAA 1286
DB 44263 CCAATTTAAAGAGTGAATTTCACTGAGAGGATCTGCAATGCTGAGAGATTTAGAA 44204
QY 1287 CAAAGTGAAG 1346
DB 44203 CAAAGTGAAG 44144
QY 1347 CCAAGTGAAG 1406
DB 44143 CCAAGTGAAG 44084
QY 1407 TATTTTAAAGATTTCAAGCTATTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAG 1466
DB 44083 TATTTTAAAGATTTCAAGCTATTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAG 44024
QY 1467 TCATGATGAGAGTGTGATTTAGACTTTTAAAGTGTGATTTTAAAGTGTGAT 1526
DB 44023 TCATGATGAGAGTGTGATTTAGACTTTTAAAGTGTGATTTTAAAGTGTGAT 43964
QY 1527 AGTGTCTGC 1536
DB 43963 AGTGTCTGC 43954

RESULT 5

US-08-965-048-6/c
; Sequence 6, Application US/08965048
; Patent No. 6323244
; GENERAL INFORMATION:
; APPLICANT: Chen, Hong
; APPLICANT: Freimer, Nelson
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF NEUROPSYCHIATRIC DISORDERS
; FILE REFERENCE: 7853-093
; CURRENT APPLICATION NUMBER: US/08/965,048
; CURRENT FILING DATE: 1997-11-05
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 45983
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-965-048-6

Query Match 40.9%; Score 627.8; DB 4; Length 45983;
Best Local Similarity 91.1%; Pred. No. 9,5e-157;
Matches 665; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 807 GTACATGATGCGACATCA CGGCGCTGAAGCCAGGTGACTTTATACACTTTGGGAGA 866
DB 44956 GAACCTTGTGATCACTCTGTGACTTTGTTCA TGGACATGAGAGCAATTAACAG 44897
QY 867 TGCACATTTTACCTGATCAATCGAGCCACTGAAAATTCAGCTTCAGCGAAACCCAG 926
DB 44896 GTGCTACATTTTGGCAAAATTAATGAGCTTATTTTGTGTTTATAGCTTCAGCGAAACCCAG 44837
QY 927 ACCCTTCCAAAGCTCAGATCTTCCGAAAAGTTGAGAAAATGATGACTTCAAACTGA 986
DB 44836 ACCCTTCCAAAGCTCAGATCTTCCGAAAAGTTGAGAAAATGATGACTTCAAACTGA 44777
QY 987 AGACTTCAATTTAGAGGCTCAATCCGATCAA CTATTAAATGGAATGGCTGTTTA 1046
DB 44776 AGACTTCAATTTAGAGGCTCAATCCGATCAA CTATTAAATGGAATGGCTGTTTA 44717
QY 1047 GGGTCTTCAAGAGAGCTGAGATTTGACCTTTAGGGGTTGGGCTGAGATGCCG 1106
DB 44716 GGGTCTTCAAGAGAGCTGAGATTTGACCTTTAGGGGTTGGGCTGAGATGCCG 44657
QY 1107 AGGTAAAGTTCTTTTGTCTTAAAGAAAGAACTAGAGTCAAAATCTGCTCCGTAC 1166
DB 44656 AGGTAAAGTTCTTTTGTCTTAAAGAAAGAACTAGAGTCAAAATCTGCTCCGTAC 44597
QY 1167 CTATCAGTATTAATTTTAAAGATGTCCTGCACTGGCAATGTACTGTGCAATCTTT 1226
DB 44596 CTATCAGTATTAATTTTAAAGATGTCCTGCACTGGCAATGTACTGTGCAATCTTT 44537
QY 1227 CCATTAATTAAGGCTTTGATTAATCTCACTGAGGCTATCTGACAACTGAGTTATGA 1286
DB 44536 CCATTAATTAAGGCTTTGATTAATCTCACTGAGGCTATCTGACAACTGAGTTATGA 44477
QY 1287 CAAAGTGAAGAAATGAATGTATGTCTTTAGAGAAATGATGAGCAATTCATC 1346
DB 44476 CAAAGTGAAGAAATGAATGTATGTCTTTAGAGAAATGATGAGCAATTCATC 44417
QY 1347 CCAAGTACTTAAGAGGTGTGATTTCAAGCTATTTTGGAAATTTTAAAG 1406
DB 44416 CCAAGTACTTAAGAGGTGTGATTTCAAGCTATTTTGGAAATTTTAAAG 44357
QY 1407 TATTTAAGATTTCAAGCTATTTCCCTCAATCTGAGGAGCTGAGTAACCATGA 1466
DB 44356 TATTTAAGATTTCAAGCTATTTCCCTCAATCTGAGGAGCTGAGTAACCATGA 44297
QY 1467 TCATGATGAGAGTGTGTTATGAATTTANAGTTGTTTATAGTTGCTATTAATGA 1526
DB 44296 TCATGATGAGAGTGTGTTATGAATTTANAGTTGTTTATAGTTGCTATTAATGA 44237
QY 1527 AGTGTCTGC 1536
DB 44236 AGTGTCTGC 44227

RESULT 6

US-09-318-448-11
; Sequence 11, Application US/09318448
; Patent No. 6210950
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Stencos, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE REFERENCE: 601-1-057
; CURRENT APPLICATION NUMBER: US/09/318,448
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 18596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-318-448-11

Query Match 39.8%; Score 611.8; DB 4; Length 18596;
Best Local Similarity 90.4%; Pred. No. 1,1e-152;
Matches 665; Conservative 0; Mismatches 65; Indels 6; Gaps 1;
QY 807 GTACATGATGCGACATCA CGGCGCTGAAGCCAGGTGACTTTATACACTTTGGGAGA 866
DB 15510 GAACCTTGTGATCACTCTGTGACTTTGTTCA TGGACATGAGAGCAATTAACAG 15569
QY 867 TGCACATTTTACCTGATCAATCGAGCCACTGAAAATTCAGCTTCAGCGAAACCCAG 926
DB 15570 GTGCTACATTTTGGCAAAATTAATGAGCTTATTTTGTGTTTATAGCTTCAGCGAAACCCAG 15629
QY 927 ACCCTTCCAAAGCTCAGATCTTCCGAAAAGTTGAGAAAATGATGACTTCAAACTGA 986
DB 15630 ACCCTTCCAAAGCTCAGATCTTCCGAAAAGTTGAGAAAATGATGACTTCAAACTGA 15689
QY 987 AGACTTCAATTTAGAGGCTCAATCCGATCAA CTATTAAATGGAATGGCTGTTTA 1046
DB 15690 AGACTTCAATTTAGAGGCTCAATCCGATCAA CTATTAAATGGAATGGCTGTTTA 15749
QY 1047 GGGTCTTCAAGAGAGCTGAGATTTGACCTTTAGGGGTTGGGCTGAGATGCCG 1106
DB 15750 GGGTCTTCAAGAGAGCTGAGATTTGACCTTTAGGGGTTGGGCTGAGATGCCG 15809
QY 1107 AGGTAAAGTTCTTTTGTCTTAAAGAAAGAACTAGAGTCAAAATCTGCTCCGTAC 1166
DB 15810 AGGTAAAGTTCTTTTGTCTTAAAGAAAGAACTAGAGTCAAAATCTGCTCCGTAC 15869
QY 1167 CTATCAGTATTAATTTTAAAGATGTCCTGCACTGGCAATGTACTGTGCAATCTTT 1226
DB 15870 CTATCAGTATTAATTTTAAAGATGTCCTGCACTGGCAATGTACTGTGCAATCTTT 15929
QY 1227 CCATTAATTAAGGCTTTGATTAATCTCACTGAGGCTATCTGACAACTGAGTTATGA 1286
DB 15930 CCATTAATTAAGGCTTTGATTAATCTCACTGAGGCTATCTGACAACTGAGTTATGA 15989
QY 1287 CAAAGTGAAGAAATGAATGTATGTCTTTAGAGAAATGATGAGCAATTCATC 1346
DB 15990 CAAAGTGAAGAAATGAATGTATGTCTTTAGAGAAATGATGAGCAATTCATC 16049
QY 1347 CCAAGTACTTAAGAGGTGTGATTTCAAGCTATTTTGGAAATTTTAAAG 1406
DB 16050 CCAAGTACTTAAGAGGTGTGATTTCAAGCTATTTTGGAAATTTTAAAG 16109
QY 1407 TATTTAAGATTTCAAGCTATTTCCCTCAATCTGAGGAGCTGAGTAACCATGA 1466
DB 16110 TATTTAAGATTTCAAGCTATTTCCCTCAATCTGAGGAGCTGAGTAACCATGA 16169
QY 1467 TCATGATGAGAGTGTGTTATGAATTTANAGTTGTTTATAGTTGCTATTAATGA 1520
DB 16170 TCATGATGAGAGTGTGTTATGAATTTANAGTTGTTTATAGTTGCTATTAATGA 16229
QY 1521 TAAAGAGTGTCTGC 1536
DB 16230 TAAAGAGTGTCTGC 16245

RESULT 7

US-09-230-637-1
; Sequence 1, Application US/09230637
; Patent No. 6264958
; GENERAL INFORMATION:
; APPLICANT: Hayward, Gary
; APPLICANT: Nicholas, John
; APPLICANT: Hardwick, J. Marie
; APPLICANT: Reitz, Marvin
; TITLE OF INVENTION: No. 6264958el Genes of Kaposi's Sarcoma
; FILE REFERENCE: 1107.78372
; CURRENT APPLICATION NUMBER: US/09/230,637
; CURRENT FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: 60/022,591

```

; PRIOR FILING DATE: 1996-07-25
; PRIOR APPLICATION NUMBER: PCT US 97/12931
; PRIOR FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-637-1

```

```

Query Match      25.2%; Score 387.4; DB 4; Length 1014;
Best Local Similarity 63.5%; Pred. No. 1.3e-93;
Matches 592; Conservative 0; Mismatches 341; Indels 0; Gaps 0;

```

```

QY 115 GCCGCTCGGAGCTGCCGCGCCGCGCCCTTGCCCCCGCGACGAGCGCGGACGCGAG 174
DB 82 GGCCTCTTGCTTGCGCCCGCGAGGTTGACCTTGTTGATTGACATGAGTTAACTGGG 141
QY 175 CGCGCTCGCGCGACGCGGAGCTGCACTCTGCGGACAGATCCAAACATCTCGCTGC 234
DB 142 GCTGAGAGCCCCCAAGAGACTTCACTGATCTCAGGCAATTGAGGAAATTTGTGCGT 201
QY 235 GCGCTCAGGAGAGAGACGACCGGACCGGACCGGACCGCTGTGATTGCGATGAGCG 294
DB 202 GGCACCGATGCACTGACCGGACCGGACCGGACCGCTGTGATTGCGATGAGCG 261
QY 295 CGCTACAGCTGAGAGATGATTCCTCTGCTGACAAACAGTGTGTTGGAAGGT 354
DB 262 AGGTATAGTCTGCGGACCACTTCCCTTACTAACCAAGGCGGTGTTTGCGAGGC 321
QY 355 GTTTTGGAGAGTCTGCTGTGTTTATCAGGAGATCCAAATGCTAAAGCTGTGCTCC 414
DB 322 GTAGTGAAGAGCTGCTTGTGTTTCTGAAGGAGATGCTGCTCAGGAGGCTTTGAGA 381
QY 415 AAGGAGATGAATCTGAGATGCGAATGATCCGAGACTTTTGGACAGCTGAGATTC 474
DB 382 AAGAGATGCAAGATGAGGACAAATAGCTCAGGAGATTCCTGCGCGGCGCGCTG 441
QY 475 TTCACGAGAGAGAGGAGCTTGCGCCCGAGTTTATGCTTCAAGTGAAGGCAATTTTGG 534
DB 442 GCGCACAGAGAGAGGAGGAGATTTGGACCTGTTTACGGTTTCAAGTGAAGGCACTTTGG 501
QY 535 GCAGATACAGAGATATGATGATGATTAATTAAGACAGGAGTTGACCACTGCAAGAGA 594
DB 502 GCGGCGTACGTGAGCGCGATGCTGATTAAGAGCCAGGCGTTTGAACATATGCTGAC 561
QY 595 GTGATTTGACCAATCAAAACCAACCTTGACGACAGAAATATCAGTACGCTTGGAT 654
DB 562 ATTGTGATTTAATTAATAAATAATCCGACAGTAAAGATCATTAATGTTGCGTGAAC 621
QY 655 CCAAGAGATTTCTCTGATGAGCGCTGCTCCATGACCAATGCGCTCCGACTGATATG 714
DB 622 CCGCGGAGCTTGTGATGAGCGCTTCCGCGCTTCACTTTGATGCAATTTTATGTA 681
QY 715 GTGAACATGAGCTGCTCTGCGACGCTGACAGAGATGCGGAGCATGCGGCTCGTGTG 774
DB 682 GCTGACGAGCTTCTCTGCTGAGCTGATCAGAGGTGCGGAGACATGCGTTGGAGATT 741
QY 775 CTTTCAACATGCGCAGCTACGCTGCTCAGTACATGATTTGCGACATCAAGGCGCTG 834
DB 742 CTTTAACTATGCGAGCTATTCCTCTTAATTAATGATTTGCTGCAATGTTATCGTCTT 801
QY 835 AAGCAGGATGCTTTATACACTTTTGGAGATGACATTAATTAACCTGAATCAATCGAG 894
DB 802 AAGACCGGAGGATTTATTAACAAGTGGAGATGCGCACTCAAAAGCATATAGAG 861
QY 895 CCACTGAATAATTCAGCTTCAAGCGAAGAACCAAGCTTTTCCAAAGCTCAGATTTCTGA 954
DB 862 CCACTACGCTGAGCTGAGCGGACCTCAACGTCCTTTCCGCGCTGAGATATCCG 921
QY 955 AAGTTGAGAAATTTGATGATTCGAAGCTGGAAGCTTTCAGATTGAAGGATCAATCCG 1014

```

```

DB 922 TCTGTTCTTCATGAGAGATTACACCTGATGATTTTGAATGCTGAGACTACTGCCG 981
QY 1015 CATCCACTATTTAAATGGAATGCGCTTTAG 1047
DB 982 CATCCACTATTCATGGAATGAGATATAG 1014

```

RESULT 8

```

US-08-770-379-17/C
; Sequence 17, Application US/08770379
; Patent No. 5849564

```

```

; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
; TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036

```

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770,379
; FILING DATE:

```

CLASSIFICATION:

```

435
ATTORNEY/AGENT INFORMATION:

```

```

NAME: White, John P.

```

```

REGISTRATION NUMBER: 28,678

```

```

REFERENCE/DOCKET NUMBER: 52342

```

```

TELECOMMUNICATION INFORMATION:

```

```

TELEPHONE: (212) 278-0400

```

```

TELEFAX: (212) 391-0525

```

```

INFORMATION FOR SEQ ID NO: 17:

```

```

SEQUENCE CHARACTERISTICS:

```

```

LENGTH: 35100 base pairs

```

```

TYPE: nucleic acid

```

```

STRANDEDNESS: double

```

```

TOPOLOGY: linear

```

```

MOLECULE TYPE: DNA (genomic)

```

```

US-08-770-379-17

```

```

Query Match      25.2%; Score 387.4; DB 2; Length 35100;
Best Local Similarity 63.5%; Pred. No. 8.2e-93;
Matches 592; Conservative 0; Mismatches 341; Indels 0; Gaps 0;

```

```

QY 115 GCCGCTCGGAGCTGCCGCGCCGCGCCCTTGCCCCCGCGACGAGCGGAGCGCGAG 174
DB 21023 GGCCTCTTGCTTGCGCCCGCGAGGTTGACCTTGTTGATTGACATGAGGTAACTGGG 20964
QY 175 CGCGCTCGCGCGACGCGGAGCTGCACTGAGGAGATCTGAGGAGATCAACATCTCCGTG 234
DB 20963 GCTGAGACGCGCCCAAGAGAACTTCAATATCTCAGGCTTGAAGGAAATTTGTGCGCT 20904
QY 235 GCGCTCAGAGAGAGAGACCGGACCGGACCGGACCGGACCTTGTGCTATTTGGGATGAGCG 294
DB 20903 GCGACGATGACTGACCGGACCGGACCGGACCGGACCGGACCTTCTCTTAATTTGGCATGAGCG 20844
QY 295 CGCTACAGCTGAGAGATGAATTCCTCTGCTGACAAACAAACGATGTTTGGAAAGGT 354
DB 20843 AGGTATGCTGCGGAGACACTTTCCTTACTAACCAAGCGGATTTTGGCGAGGC 20784
QY 355 GTTTTGAAGAGTGTGCTGTTTATCAAGGATCCAAATGCTTAAGAGCTGTCTTCC 414

```

Db	20783	GTAGTGCAGAGCTGCTTTTGTTTGTTGAAAGGGAGATCTGATCTCCAGGGACTTTTCAAG	20724
QY	415	AAGGAGTGAATCTGGGATGCCATGAGATCCGAGACTTTTGAACGCTTGGGATTC	474
Db	20723	ACAGAGTCAAGATATAGGACAAAATATGCTCCAGGAGTTCTGTGACGGGCGCGACTG	20664
QY	475	TCCACCAAGAAGAGGGGACTTTGGGCCCAATTTTATGCGTTCAGATGGAGCATTTTGGG	534
Db	20663	GGGACACAAAGGAGGGGGATTGGGACCTGTATACGTTTCCAGTGGAGGCACTTGGG	20604
QY	535	GCAGATACAGAGTATGAGATCAATATTTCAAGACAGGGAGTTGACCAACTGCAAG	594
Db	20603	GGGGGTACGTGGACGGGATGTCTGATCTATACAGGCCAGGGGTTTGCACATTTGCTGAC	20544
QY	595	GTGATTCACACCATCAAAACCAACCTGACAGACAAAGATCATCATGTGGCTTGGAT	654
Db	20543	ATTGTGGATTATATATATATATATCCGACAGATGAGGATCATATGTGTGGTGGAC	20484
QY	655	CCAAAGATCTCTCTGTGATGGGGCTGCTCCATGCGATGACCTCTGGCACTTGTATGG	714
Db	20483	CCGGGGGACTGTGTGTGAATGGGCTTCGGCTGTGACTTGTATATTTATATATATAT	20424
QY	715	GTGAACAGTAGAGCTGTCTCTGACGCTGACACAGATCCGGAGACATGGGCTCGGTGG	774
Db	20423	GCTGACGTGAGCTTTCTGTCACTGTATAGAGGTCCGGAGACATGGGTTTGGAGTT	20364
QY	775	CCTTTCAACATGCGACGTAAGCCCTGTCTCAGTACATGATTTGCGACATACGGGCTTG	834
Db	20363	CGTTTACATGCGACACTTTCCTCTTAATATATATGCTTGTCTCATATGTATACGTCTT	20304
QY	835	AAGCCAGGTGACTTTATACACATTTGGGAGATGCAATTTTACCTGATACATCGAG	894
Db	20303	AGACCGGGAGTTTATTCACAGTTGGGAGATGCCACATCTCAAAACGCAATATGAG	20244
QY	895	CCACTGAATATTCAGCTTTCAGCAGAGACCCAGACCTTTCCAAAGCTCAGATTTCTTCA	954
Db	20243	CCACTAGGCTGCACTGAGCGGCACTCAGTCCCTTTCCGGCCTTGAATATCTCGG	20184
QY	955	AAAGTTGAGAAATTTGATGACTTCAAGCTGAGATCTTTTCAATGTAAGGGTACATTCG	1014
Db	20183	TCTGTTCTTCCATGAGAGGTTTACACTGATGATTTTGACTGTGTGACTACTGCGCG	20124
QY	1015	CATCCACATTTAAATGAGAAATGGCTGTTG	1047
Db	20123	CATCTACCATTCGTATGGAGAAATGGCAGTATAG	20091

Db 20303 AGACCCGGGAGTTTATTCACACCTTGGGAGATGCCACCTCTCAAAAACGCATATAGAG 20244
Qy 895 CCACGTGAATTCAGCTTCAGCGAGAACCCAGACCTTTCCGAAAGCTCAGATTTCTTCA 954
Db 20243 CCACGTGAATTCAGCTTCAGCGAGAACCCAGACCTTTCCGAAAGCTCAGATTTCTTCA 20184
Qy 955 AAGGTGAGAAATTCAGCTTCAGCGAGAACCCAGACCTTTCCGAAAGCTCAGATTTCTTCA 1014
Db 20183 TCTGTTCTTCATGAGAGATTTACACCTGATGATTTAGACGTGTGACCTCTGCGG 20124
Qy 1015 CATCCACTATTAATGAGAAATGCGCTTTAG 1047
Db 20123 CATCCACTATTCATGATGAGAAATGCGCTTTAG 20091

RESULT 10
US-09-230-371A-17/c
; Sequence 17, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; FILE REFERENCE: 45185-G-PC-T-17
; CURRENT APPLICATION NUMBER: US/09/230,371A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 17
; LENGTH: 35100
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-371A-17

Query Match 25.2%; Score 387.4; DB 4; Length 35100;
Best Local Similarity 63.5%; Pred. No. 8.2e-93;
Matches 592; Conservative 0; Mismatches 341; Indels 0; Gaps 0;
Qy 115 GCGGCTCGAGAGCTGCCCGCGCCCTTGGCCCGCCGCGACGAGGAGCGGACGCCAG 174
Db 21023 GCGGCTCGAGAGCTGCCCGCGCCCTTGGCCCGCCGCGACGAGGAGCGGACGCCAG 20964
Qy 175 CCGCGTCCGCGCAGCGGAGAGCTGAGTACCTGCGGAGAGATCCAAACATCTCCGCTGC 234
Db 20963 GCTGAGACCGCCCGCAGAGAGATTCAGTATCTCAGGAGTGAAGGAAATTTGTGCCGT 20904
Qy 235 GCGGCTCGAGAGAGCGACCGCGACCGGACCGGACCTGTGCGATTTGGCATGAGAGG 294
Db 20903 GCGGCTCGAGAGAGCGACCGCGACCGGACCGGACCTGTGCGATTTGGCATGAGAGG 20844
Qy 295 CGCTACAGCTTGAGAGATGAATTCCTCTGCTGACACCAAAAGCTGTCTTGGAAAGGT 354
Db 20843 AGGTATAGCTGCGGAGACCTTTCCCTTACTAACCAAAAGCGGTGTTTGGCAGAGGC 20784
Qy 355 GTTTTGGAGAGATGCTGTGCTGTTATTCAGAGGATCCAGAAATGCTAAAGAGCTGTCTTC 414
Db 20783 GTTGTGCAAGAGCTGTGCTGTTTCTGAAGGAGAGTACTGACTCCAGGAGCTTTCAAGA 20724
Qy 415 AAGGAGTGAATTCGAGATGCAATGATCCGAGACTTTTGGACACCTCGGATTC 474
Db 20723 ACGGAGTGAATTCGAGATGCAATGATCCGAGACTTTTGGACACCTCGGATTC 20664
Qy 475 TCCACGAGAGAGAGGAGCTTGGGCGCAGTTTATGCTTCAGTGGAGGCAATTTTGGG 534
Db 20663 GCGCAGAGAGAGGAGGAGATTTGGGCGCAGTTTATGCTTCAGTGGAGGCAATTTTGGG 20604

Qy 535 GCAGATTAACAGAGATTAAGAGATCAATTAATTCAGACAGGAGATTGACCACTGCAAGA 594
Db 20603 GCGGCTGACGTGAGACCGGAGATTCGATATACAGGAGGAGTTGACCAATTCGTAC 20544
Qy 595 GTGATTAACAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAG 654
Db 20543 ATTTGATTAATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAG 20484
Qy 655 CCAAGATTAATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAG 714
Db 20483 CCGGCGAGCTTGTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 20424
Qy 715 GTGAGAGTGAAGCTGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 774
Db 20423 GCTGAGAGTGAAGCTGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 20364
Qy 775 CTTTCAACATGCGCAGCTACGCGCTGCTCAGTACATGATGATGCGACATCAGGAGCTTG 834
Db 20363 CTTTCAACATGCGCAGCTACGCGCTGCTCAGTACATGATGATGCGACATCAGGAGCTTG 20304
Qy 835 AAGCGAGTGAATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAG 894
Db 20303 AGACCCGGGAGATTAATTCACAGGAGAGATTCACATCTACAAACGATATAGAG 20244
Qy 895 CCACGTGAATTCAGCTTCAGCGAGAACCCAGACCTTTCCGAAAGCTCAGATTTCTTCA 954
Db 20243 CCACGTGAATTCAGCTTCAGCGAGAACCCAGACCTTTCCGAAAGCTCAGATTTCTTCA 20184
Qy 955 AAGGTGAGAAATTCAGCTTCAGCGAGAACCCAGACCTTTCCGAAAGCTCAGATTTCTTCA 1014
Db 20183 TCTGTTCTTCATGAGAGATTTACACCTGATGATTTAGATGATGATGATGATGATGAT 20124
Qy 1015 CATCCACTATTAATGAGAAATGCGCTTTAG 1047
Db 20123 CATCCACTATTCATGATGAGAAATGCGCTTTAG 20091

RESULT 11
US-09-203-895-2
; Sequence 2, Application US/09203895
; Patent No. 6436410
; GENERAL INFORMATION:
; APPLICANT: Krishnan, B. Rajendra
; APPLICANT: Yoder, S. Christine
; APPLICANT: Durschi, Becky A.
; TITLE OF INVENTION: DNA ENCODING NEOSPORA DIHYDROFOLATE REDUCTASE
; FILE REFERENCE: PC9890A
; CURRENT APPLICATION NUMBER: US/09/203,895
; CURRENT FILING DATE: 1998-12-02
; EARLIER APPLICATION NUMBER: 60/067,507
; EARLIER FILING DATE: 1997-12-04
; EARLIER APPLICATION NUMBER: 60/095,213
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentln Ver. 2.0 - beta
; SEQ ID NO 2
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: Neospora caninum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1839)
US-09-203-895-2

Query Match 21.7%; Score 333.4; DB 4; Length 1839;
Best Local Similarity 62.9%; Pred. No. 4.3e-79;
Matches 552; Conservative 0; Mismatches 316; Indels 9; Gaps 2;
Qy 180 TCCGCGCAGCGGAGAGCTGCAATGATCCGAGAGATTCACAAATCCTCGCTGCGGCT 239
Db 963 TCCGCGCAGCGGAGAGATTCAGATGATCCGAGAGATTCACAAATCCTCGCTGCGGCT 1022

Oy	240	CAGAAAGGACACCCGACCGGAGACCGGACCCCTGGTCGATATCGAGATCGAGCGCCCTA	299
Db	1023	GACAAATGACGACCGAACCGGCGTTTGAGATCATCTTCAAAGTTGGCGCTGACCAATGCGGTT	1081
Oy	300	CAGCCTGAGAGATGAATTCCTCTCTCTGACCAACAAACGTGTGTTCTGGAAAGGTTGTTT	359
Db	1083	CTCGCTGATPAAAGCCTTCTCCCTCTCTCAACAAAGCGTGTGTTCTGGAAAGGATCTT	1144
Oy	360	GGAGAGTTGCTGTGCTTTATCAAGGATTCACAATATCTAAAGACTGTCTTCCAAAGG	419
Db	1143	CGAGGAGCTGTGTGGTTCTATCCGCGTGAACAGAACCAATCACTTCTTGAAAAGG	1202
Oy	420	AGTGAATCTGGGATGGCAATGATCCGACACTTTTGGACAGCCTGGATTCCTCAC	479
Db	1203	CGTAAGATCTGGGACAAAGATGTGAACAAAGAGTTCTTGTATTCACGAATCTTTCCCA	1262
Oy	480	CAGAGAAAGAGGAGTCTGGGCCAGTTTATGGCTTCCATGGAGGCAATTTTGGGACAG	539
Db	1263	CGAGAGGTGAGAGCATCGGCGCGGGTTACGGCTTCAATGGAGACAATTGGCGCGAC	1322
Oy	540	ATPCAGATATTTGGATTCAGATTTTATAGGACAGGAGTTGACCACTGSCAAAGATGAT	599
Db	1323	CTACAGGACATGACACAGCATCACATGCTGAGGCGCTAACCACTGAAGAGGTGAT	1382
Oy	600	TGACACCATCAAAACCAACCCCTGACGACAAAGAAATCAATGTGGCTTGGAAATCCAA	659
Db	1383	CAACATGTGTGAACCAATCCAAACAGACCGGCGATGTCAATACCGTCTGGAAACCTGG	1442
Oy	660	AGATCTTCTCTGATGGCGCTGCTCCATGACCAATGCTCCAGTTCTATGTGGTA	718
Db	1443	GGCGCTGACGAATAGCGCTTGGCGCTTGCACCTTGTGTGACATTTCAATGGAGAA	1502
Oy	719	--ACGTGAGCTCTCTGCGACGTGTACAGAGATCGGAGAGCAATGGCGTCGTGGCGC	776
Db	1503	CGACGAGACTTGTCTTGTGTCAATGATACAGCGGTCTTGGAGAGTTGGGCGTGGCC	1562
Oy	777	TTTCAACATGSCCAGCTACGCCCTGTCTCAGATCAATGATTTGCGACATCACGGGCTGAA	836
Db	1563	GTTCAACATTTGCGTCTATTCCTCTTTCAGACGCTCATGGTTGCGGACGTGTGAACCTGAA	1622
Oy	837	GCGAGTGACTTTATACACTTTTGGAGATGCAATTTTACTGATATCATCTGAGCC	896
Db	1623	GCGGAGAGAGTTCACTTCAATGGGCAACAGCAGCTTCACTCAACCAAGTCTGAGG	1682
Oy	897	ACGGAATTTAGCTTCAACGAGAACCCAAACCTTCCCAAGAGCTGAGATTCCTCAAA	956
Db	1683	CTTGAGGAGAGCTGCGAGAGAACCGAATCTTTCCCATGTGAACATCTTGAAACA	1742
Oy	957	AG-----TTGAGAAATTTGATGACTTCAAGCTGAGACTTTCAGATTGAAGGTTACAA	1011
Db	1743	GGAACGATCCAGAAATGACAGCTTCAACCGCGCAGAGATTCGAGGTGTGTGGCTACGT	1802
Oy	1011	TGCGATCCAACTTAAATAGAAATGGCTGTGTTG	1047
Db	1803	GCGCATGTGAATCCAGATCCAGATGAGATGGCTGTGTTG	1839

RESULT 12
 US-08-965-048-1/c
 ; Sequence 1, Application US/08965048
 ; Patent No. 6323244
 ; GENERAL INFORMATION:
 ; APPLICANT: Chen, Hong
 ; APPLICANT: Frelmer, Nelson
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
 ; TITLE OF INVENTION: TREATMENT OF NEUROPSYCHIATRIC DISORDERS
 ; FILE REFERENCE: 7853-093
 ; CURRENT APPLICATION NUMBER: US/08/965,048
 ; CURRENT FILING DATE: 1997-11-05
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1817

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)..(1149)
;
US-08-965-048-1

```

Query Match	14.5%;	Score 222.8;	DB 4;	Length 1817;
Best Local Similarity	76.4%;	Pred. No. 1.2e-49;		
Matches 272; Conservative	0;	Mismatches 84;	Indels 0;	Gaps 0;

Qy	807	GCATGATATGGCGCAATCAACG393CCTGAAGCGA3GTGATATATACCACTTTGG3A3A	866
Db	1616	GAACCTGTTTGATCAACATCCCTG9TAATTGTTTATG3AATAGAGCAATTTCACACAG	1557
Qy	867	TGCACATATTACCTGATATCAATCCAGCCACTGAAATTCAGCTTCAGCGAACCAG	926
Db	1556	GTCGTAACATTTATGCGAAATAATG6CTTATTTTGTTTTACCTTCAGCGAACCAG	1497
Qy	927	ACCTTCCAAAGCTCAGATCTTCGAAAAGTTAGAAAATGTGACTTCACAACTGA	986
Db	1496	ACCTTCCAAAGCTCAGATCTTCGAAAAGTTAGAAAATGTGACTTCACAACTGA	1437
Qy	987	AACCTTCAGATGTAAGG3TCAATCCGATCAACATTTTAAATAGAAATG3CTGTTA	1046
Db	1436	AACCTTCAGATGTAAGG3TCAATCCGATCAACATTTTAAATAGAAATG3CTGTTA	1377
Qy	1047	GG3T3CTTTTAAAGAGCTGTAAGATATATGCACTTTTAGGG3TGGG3CTGAATGCCG	1106
Db	1376	GG3T3CTTTTAAAGAGCTGTAAGATATATGCACTTTTAGGG3TGGG3CTGAATGCCG	1317
Qy	1107	AG3TAAAGTCTTTT3CTCTAAAGAGAAAGAACTAG3TCAAAATCTGATCCG	1162
Db	1316	AG3TAAAGTCTTTT3CTCTTAAAGATTCCTAGAG3TGAATCTTCTTCTAGCTG	1261

RESULT 13

US-09-221-01/B-1092/C
; Sequence 1092, Application US/09221017E
; Patent No. 6444799
GENERAL INFORMATION.

GENERAL INFORMATION:
APPLICANT: ROSS, BRUCE C.
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SENTENCES: 1120

NUMBER OF SEQUENCES: 1120
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET, 755 PACE MTLI BOAD

SINCE: 1990
CITY: Palo Alto
STATE: CA
COUNTRY: USA

ZIP: 94304-1018
COMPUTER READABLE
MEDIUM TYPE: Diskette

```

;
;
;      COMPUTER : IBM Compatible
;      OPERATING SYSTEM : Windows
;      SOFTWARE : FastSEO for Windows Version 2.0b

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER:  US/09/221,017B
FILING DATE: 23-DEC-1998

```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1182

FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1546

; FILING DATE: 30-JAN-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PP2911

;; FILING DATE: 09-APR-1998
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/AU98/01023

FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:


```

; NAME: MONROY, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 1092:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7685 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; MOLECULE TYPE: circular
; TOPOLOGY: circular
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1...7685
; US-09-221-017B-1092

Query Match      10.4%; Score 160.2; DB 4; Length 7685;
Best Local Similarity 52.7%; Pred. No. 1,2e-32;
Matches 443; Conservative 0; Mismatches 338; Indels 60; Gaps 2;

QY 186 GCACGGGAGAGCTGCATGCTGGGGGACATCCACACATCCCGCTGGCGGTACAGAA 245
DB 4571 GCGTGTATATAACATGATTTGGATCTCTTCGCCGATCTTGAGAGGGAGTACGAA 4512
QY 246 GGAAGACCGGACGGGACCGACCCCTGTGCTATTGCGGATGCGAGCGCTACAGCT 305
DB 4511 AGAAGACCGGACCGGAGTACGATCATGCGGTATTCGCGCATCAGATGTGCTTGA 4452
QY 306 GAGAGATGAATTCCTCTCTCTGTCGACCAACCAACGTGTCTGGAAGGGTGTGGA 365
DB 4451 TGAAGAGGGATTCCTCTCTCTGTCGACCAACCAACGTGTCTGGAAGGGTGTGGA 4392
QY 366 GTTGTGTGTGTTATCAAGAGATCCAAATGCTAAAGAGCTGTCTTCCAAAGGAGTGA 425
DB 4391 ACTGTTGTGTTCTCTCAAGAGTGAACAGACGATTCATCTCAGAAACATGAGGTAC 4332
QY 426 AATCTGGAGATCCATGATGCCAGACTTTTGGAGAGCTGTGCTCCACCAAGA 485
DB 4331 CATTTGAAAGG-----AAAGGCTGATGA 4308
QY 486 AGAAGGGGACTTGGGCCAGTTTATGCTTCAGTGGAGGCAATTTTGGGGCAGATACG 545
DB 4307 GAAAGGGGATTTGGGACCTATATGCTTACAGTGGGCGAGTTGGCG-----4259
QY 546 AGATATGGAATCAAGATTATTCAGACAGGAGTTGACCAATGCAAGAGTGAATGAC 605
DB 4258 -----GACTATGGGGAGGGGATATGACCAATGCAAGAGTGAATGAGGCT 4212
QY 606 CATCAAAACCAACCTTGAACAGAGAAATCAATGAGTGGCTGTGATTCAGAGATCT 665
DB 4211 GGTGGCGAAGAACCGGACCGCGGCAATATGATGCTCTGGAAATGATGAGGCT 4152
QY 666 TCTCTGATGAGCGTGTCTTCATGCTTCCATGCTTGGCAATTTTATGATGTAACAG 725
DB 4151 TCCCATATGATGATGCTTCCATGCTTCCATGCTTGGCAATTTTATGATGTAACAG 4092
QY 726 GGTGCTGCGACGCTGACAGAGATCGGAGACATGAGGCTGTGATCTCTTCAACAT 785
DB 4091 GGTGAGCTGCACTCTATAGGAGATGCGAATTTCTTCTGGGATACCTTGAAT 4032
QY 786 CGCCAGCTAGCCCTGCTCACTGATGATGCGACATCAAGGCTGTAAGCCAGTGA 845
DB 4031 AGCTTCGATGCACTCTCTGTAACAATATGCTCATGTCACAGGGGTGAACCGGGCG 3972
QY 846 CTTATATACACTTTGGGAGATGCAATTTTACCTGATATCAATGAGGCACTGAAT 905

```

```

DB 3971 ATTGTACTATACATGCGGAGCGCCCATATCTATCTTAACACTCCAGAGGTACAGCT 3912
QY 906 TCGGCTTCAAGCAGAGAACCCAGACCTTCCCAAGCTCAGATCTTTCGAAAGTTGAGAA 965
DB 3911 ACAACTGAGTCTGAGCCGAGAGGCTTGCCCAATGATGATGATTCGAGGTGCCA 3852
QY 966 AATTGATGACTTCAAGCTGAGAGCTTTCAGATTGAGAGGTACATTCGATCCAACTAT 1025
DB 3851 TCTCTTGCATCTCCATACAGAGACTTCCATCGAAGGATGATGCTCGATCCGACAT 3792
QY 1026 T 1026
DB 3791 T 3791

RESULT 14
US-09-203-895-1
; Sequence 1, Application US/09203895
; Patent No. 6436410
; GENERAL INFORMATION:
; APPLICANT: Krishnan, B. Rajendra
; APPLICANT: Yoder, S. Christine
; APPLICANT: Durtschi, Becky A.
; TITLE OR INVENTION: DNA ENCODING NEOSPORA DIHYDROFOLATE REDUCTASE
; TITLE OR INVENTION: THYMIDYLATE SYNTHASE
; FILE REFERENCE: PC9690A
; CURRENT APPLICATION NUMBER: US/09/203,895
; EARLIER FILING DATE: 1998-12-02
; EARLIER APPLICATION NUMBER: 60/067,507
; EARLIER FILING DATE: 1997-12-04
; EARLIER APPLICATION NUMBER: 60/095,213
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO: 1
; LENGTH: 9603
; TYPE: DNA
; ORGANISM: Neospora caninum
; FEATURE:
; NAME/KEY: gene
; LOCATION: (2405)..(8199)
; US-09-203-895-1

Query Match      8.1%; Score 124.2; DB 4; Length 9603;
Best Local Similarity 66.8%; Pred. No. 5e-23;
Matches 193; Conservative 0; Mismatches 93; Indels 3; Gaps 1;

QY 673 ATGCGCTGCTCTCATGCAATGCTCTTGCAGATCTTATGCTGA---ACAGTGAAGCTG 729
DB 7275 ATGCGCTGCTCTCTCATGCAATGCTCTTGCAGATCTTATGCTGA---ACAGTGAAGCTG 7334
QY 730 TCTTCCAGCTGATACAGAGATCGGAGACATGCGGCTCGGTGCTTCAACATGCGC 789
DB 7335 TCTTCCAGCTGATACAGAGATCGGAGACATGCGGCTCGGTGCTTCAACATGCGC 7394
QY 790 AGCTAGCGCGCTCTCATGCAATGCTCTTGCAGATCTTATGCTGA---ACAGTGAAGCTG 849
DB 7395 TCTTCCAGCTGATACAGAGATCGGAGACATGCGGCTCGGTGCTTCAACATGCGC 7454
QY 850 ATACACACTTGGAGATGACATATTTACTGATACATCGAGCACTGAAATTCAG 909
DB 7455 ATTCACCTTCATGAGCAACAGACGCTTATCTGAAACACAGCTGAGGCTCTGAAGAGAG 7514
QY 910 CTTCAAGAGAAACCCAGACCTTCCCAAGCTCAGATCTTCTGAAAG 958
DB 7515 CTGGCAGAGAAACCCAGACCTTCCCAAGCTCAGATCTTCTGAAAG 7563

RESULT 15
US-08-714-918-7/c
; Sequence 7, Application US/08714918
; Patent No. 6037123
; GENERAL INFORMATION:

```


APPLICANT: Benton, Bret
 APPLICANT: Lee, Ying
 APPLICANT: Malouin, Francois
 APPLICANT: Martin, Patrick K.
 APPLICANT: Schmidt, Molly B.
 APPLICANT: Sun, Dongxu
 TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
 TITLE OF INVENTION: TARGET GENES
 NUMBER OF SEQUENCES: 111
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 STREET: Suite 4700
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071-2066
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 MEDIUM TYPE: Storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/714,918
 FILING DATE: September 13, 1996
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/009,102
 FILING DATE: December 22, 1995
 APPLICATION NUMBER: 60/003,798
 FILING DATE: September 15, 1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Warburg, Richard J.
 REGISTRATION NUMBER: 32,327
 REFERENCE/DOCKET NUMBER: 222/005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEEX: 67-3510
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3479 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

QY	1	GGGGGGGGGGGAAACAATTGGCTGTGCTCTTCGTCGCGCGCACTTGGCTGCTCCGCT	60
Db	1	GGGGGGGGGGGAAACAATTGGCTGTGCTCTTCGTCGCGCGCACTTGGCTGCTCCGCT	60
QY	61	CCCGCGCGCACTTGGCTGTGCTCTTCGTCGCGCGCGCGCAATGCTGTGGCGCGC	120
Db	61	CCCGCGCGCACTTGGCTGTGCTCTTCGTCGCGCGCGCGCAATGCTGTGGCGCGC	120
QY	121	TGGAGCTGTGCGCGCGCGCCCTTGGCCCCCGCGCAAGAGCGGGAGCGCGAGCGCGT	180
Db	121	TGGAGCTGTGCGCGCGCGCCCTTGGCCCCCGCGCAAGAGCGGGAGCGCGAGCGCGT	180
QY	181	CCGCGCGCACGGGAGTGTGCAATCCTTGGGCGAATCCAAACAATCCTCGCTGCGGCTC	240
Db	181	CCGCGCGCACGGGAGTGTGCAATCCTTGGGCGAATCCTCGCTGCGGCTC	240
QY	241	AAGAAAGCAACGCAACGCGGACACCGGACCCTGTGGATTTCCGCAATGCAAGCGCGCTAC	300
Db	241	AAGAAAGCAACGCAACGCGGACACCGGACCCTGTGGATTTCCGCAATGCAAGCGCGCTAC	300
QY	301	AGCTGAGAGATGATTCCTCTGTGTCGAACCAAAAGCTGTTCGGAAGGAGTCTTTG	360
Db	301	AGCTGAGAGATGATTCCTCTGTGTCGAACCAAAAGCTGTTCGGAAGGAGTCTTTG	360
QY	361	GAGGAGTTGCTGTGCTTTTCAAGGATCCAAATGCTAAAGGCTGTCTTCAAGGGA	420
Db	361	GAGGAGTTGCTGTGCTTTTCAAGGATCCAAATGCTAAAGGCTGTCTTCAAGGGA	420
QY	421	GTGAATAATCTGGGATCCGAATCCCGAGCTTTTGGACAAGCTGGGATTTTCCACC	480
Db	421	GTGAATAATCTGGGATCCGAATCCCGAGCTTTTGGACAAGCTGGGATTTTCCACC	480
QY	481	AGGAAAGAGGGGCTTGGGCCAGTTTAAAGCTTCCAGTGTGAAGGCAATTTTGGGCAAA	540
Db	481	AGGAAAGAGGGGCTTGGGCCAGTTTAAAGCTTCCAGTGTGAAGGCAATTTTGGGCAAA	540
QY	541	TACAGAATATGGAATCAATATTTCAAGACAAGGAGTTTGAACAATGCAAAAGATGATTT	6000
Db	541	TACAGAATATGGAATCAATATTTCAAGACAAGGAGTTTGAACAATGCAAAAGATGATTT	6000

OY	601	GACACCATCAAAACCAACCCCTGACAGCAGAAGATCATATGTGGCTTSGAATCCAAGA	660
Db	601	GACACCATCAAAACCAACCCCTGACAGCAGAAGATCATATGTGGCTTSGAATCCAAGA	660
OY	661	GATCTTCCTCTGAATGGGCGTGCTCCAGGCCCATGCCCTCTGCAGATTCTATGTGTGAAC	720
Db	661	GATCTTCCTCTGAATGGGCGTGCTCCAGGCCCATGCCCTCTGCAGATTCTATGTGTGAAC	720
OY	721	AGTAGAGCTGTCCTCGACAGACTGTACAGAGATCGSAGAACATGGGCGCTCGGTGTGCTTTTC	780
Db	721	AGTAGAGCTGTCCTCGACAGACTGTACAGAGATCGSAGAACATGGGCGCTCGGTGTGCTTTTC	780
OY	781	AACATCGCAGACTAGCCCTTGCTACGTATCATGTATGGCACATCA CGGACCTGAAGCCA	840
Db	781	AACATCGCAGACTAGCCCTTGCTACGTATCATGTATGGCACATCA CGGACCTGAAGCCA	840
OY	841	GGTAGCTTTATACACACTTTGGAGATGACAATTTTACTGAATACATTCGAGCCACTG	900
Db	841	GGTAGCTTTATACACACTTTGGAGATGACAATTTTACTGAATACATTCGAGCCACTG	900
OY	901	AAAATTAAGCTTTAGCGAGAACCCBAACCTTTCCCAAAGTCGAGATCTTTCGAAAAGTT	960
Db	901	AAAATTAAGCTTTAGCGAGAACCCBAACCTTTCCCAAAGTCGAGATCTTTCGAAAAGTT	960
OY	961	GAGAAATTTGATGACCTTGAAAGCTGGAAGACTTTGAGATTGAAGGGTACATTCGCATCCA	1020
Db	961	GAGAAATTTGATGACCTTGAAAGCTGGAAGACTTTGAGATTGAAGGGTACATTCGCATCCA	1020
OY	1021	ACTATTTAAAAATGGAAATGGCTGTTTTAGGGTGTCTTCAAGAGACCTNGAAGGATTTGTCA	1080
Db	1021	ACTATTTAAAAATGGAAATGGCTGTTTTAGGGTGTCTTCAAGAGACCTNGAAGGATTTGTCA	1080
OY	1081	GCTCTTAGGGGTGTGGGCTGATGATGCCAGGGTAAAGTTCTTTTGTCTTAAAGANAAGG	1140
Db	1081	GCTCTTAGGGGTGTGGGCTGATGATGCCAGGGTAAAGTTCTTTTGTCTTAAAGANAAGG	1140
OY	1141	AAC TAGCTCAAAAATCTGTCGGTAGCCTATCA GTTATTAATTTTAAAGATGTGGCACT	1200
Db	1141	AAC TAGCTCAAAAATCTGTCGGTAGCCTATCA GTTATTAATTTTAAAGATGTGGCACT	1200
OY	1201	GGCAAAATGTAACGTGCGAAGTCTTTCACAAAATAAAGCTTTGAGTTAACTCACTGAAGG	1260
Db	1201	GGCAAAATGTAACGTGCGAAGTCTTTCACAAAATAAAGCTTTGAGTTAACTCACTGAAGG	1260
OY	1261	GTATCTGACATCTGAGGTTATGACAAAGTGAAGAGAAATGAATGTATGTGCTCTTAG	1320
Db	1261	GTATCTGACATCTGAGGTTATGACAAAGTGAAGAGAAATGTATGTGCTCTTAG	1320
OY	1321	CAAAAACATGTATGTGCATTTCAATCCCGATCACTTAAAGAGGTGTGATTTTCAG	1380
Db	1321	CAAAAACATGTATGTGCATTTCAATCCCGATCACTTAAAGAGGTGTGATTTTCAG	1380
OY	1381	AAGCTATTTTGGAAATATTTTAAATATTTTAAAGATTTTCA CAAGCTATTCCTCAAAAT	1440
Db	1381	AAGCTATTTTGGAAATATTTTAAATATTTTAAAGATTTTCA CAAGCTATTCCTCAAAAT	1440
OY	1441	CTGAGGGAGCTGAGTAACACATGCATGCATGTATGAGAGTGTGTTATTAACCTTAAAGT	1500
Db	1441	CTGAGGGAGCTGAGTAACACATGCATGCATGTATGAGAGTGTGTTATTAACCTTAAAGT	1500
OY	1501	TGTTTTATATGTTGCTATTAATAAAGAGTGTCTGAC	1536
Db	1501	TGTTTTATATGTTGCTATTAATAAAGAGTGTCTGAC	1536
RESULT 2			
AAS94945			
ID	AAS94945 standard; DNA; 3298 BP.		
XX			
XX	AAS94945;		
DT	14-FEB-2002 (first entry)		

XX Human DNA sequence #200 expressed during foam cell differentiation.
 DE Human: foam cell differentiation; atherosclerosis; cerebral stroke;
 XX Human: foam cell differentiation; atherosclerosis; cerebral stroke;
 KM Cardiovascular disorder; coronary artery disease; gene therapy; ds.
 XX Hemo sapiens.
 OS Hemo sapiens.
 XX MO200177389-A2.
 PN 18-OCT-2001.
 PD 04-APR-2001; 2001MO-US11128.
 PF 05-APR-2000; 2000US-195106P.
 PR (INCY-) INCYTE GENOMICS INC.
 PA Shiffman D, Somogyi R, Lawn R, Seilhamer JF, Porter GJ, Mikita T,
 P1 Tai J;
 P1 WPI; 2002-010925/01.
 DR Composition useful for diagnosis of conditions, disorders or diseases
 XX associated with atherosclerosis, comprises several polynucleotides that
 PT are differentially expressed in foam cell development -
 XX Claim 1; Page 250-251, 315pp; English.
 PS The present invention relates to the isolation of human polynucleotide
 CC sequences that are differentially expressed during foam cell
 CC differentiation. The polynucleotide sequences of the invention or a
 CC composition comprising these polynucleotides are useful as a high
 CC throughput method for detecting altered expression of one or more
 CC polynucleotides in a sample. The polynucleotides can be used in the
 CC diagnosis of disorders associated with foam cell development such as
 CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
 CC coronary artery disease. The polynucleotide sequences can also be used
 CC as PCR primers and probes. The polynucleotides of the invention are also
 CC useful in gene therapy. AAS94/46-AAS95021 represent the human
 CC polynucleotide sequences of the invention which are differentially
 CC expressed during foam cell differentiation.
 CC Sequence 3298 BP; 898 A; 709 C; 744 G; 919 T; 28 other;
 SO Query Match 97.2%; Score 1493.4; DB 24; Length 3298;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1510; Conservative 0; Mismatches 5; Indels 6; Gaps 1;

QY 362 AAGGATCCACAAAGCTGTAAGAGCTGCTCCAGAGAGAGTAAATCTGGATGCCAAT 441
 DB 361 AAGGATCCACAAAGCTGTAAGAGCTGCTCCAGAGAGAGTAAATCTGGATGCCAAT 420
 QY 442 GGAATCCGAGACTTTTGGACAGCTTGGGATTTCTCCACGAGAGAGAGGGAATTGGGCT 501
 DB 421 GGATCCGAGACTTTTGGACAGCTTGGGATTTCTCCACGAGAGAGAGGGAATTGGGCT 480
 QY 502 CCAATTATGAGCTTCAGAGAGGAGCATTTTGGGAGAGATTCAGAGATATGAAATCAGAT 561
 DB 481 CAGATTATGAGCTTCAGAGAGGAGCATTTTGGGAGAGATTCAGAGATATGAAATCAGAT 540
 QY 562 TATTCAGAGACAGGAGATTGACCAACTGCAAGAGTATGACCACTATCAAAACCAACCT 621
 DB 541 TATTCAGAGACAGGAGATTGACCAACTGCAAGAGTATGACCACTATCAAAACCAACCT 600
 QY 622 GACGACAGAGAAATATATATATGAGCTTGGGATTCAGAGATCTTCTCTGATGAGGCTG 681
 DB 601 GACGACAGAGAAATATATATATGAGCTTGGGATTCAGAGATCTTCTCTGATGAGGCTG 660
 QY 682 CTTCCAGGCAATGAGCTTCCAGAGTCTTATATGAGTGAACATGAGCTTCTCCAGAGCTG 741
 DB 661 CTTCCAGGCAATGAGCTTCCAGAGTCTTATATGAGTGAACATGAGCTTCTCCAGAGCTG 720
 QY 742 TACCAAGATCGGAGACATGAGGCTGAGCTTCAACATGAGCTTCAACATGAGGCTG 801
 DB 721 TACCAAGATCGGAGACATGAGGCTGAGCTTCAACATGAGCTTCAACATGAGGCTG 780
 QY 802 CTACAGATGATGATGAGCAATGAGGCTGAGGCTGAGGCTGAGGCTGATATACACTTGG 861
 DB 781 CTACAGATGATGATGAGCAATGAGGCTGAGGCTGAGGCTGAGGCTGATATACACTTGG 840
 QY 862 GAGAGATGACATATATATGAGTATGACATGAGGCTGAGGCTGAGGCTGAGGCTG 921
 DB 841 GAGAGATGACATATATATGAGTATGACATGAGGCTGAGGCTGAGGCTGAGGCTG 900
 QY 922 CCGAGAGCTTTCAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 981
 DB 901 CCGAGAGCTTTCAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 960
 QY 982 GCTGAAGACTTTCAGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1041
 DB 961 GCTGAAGACTTTCAGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1020
 QY 1042 GTTAAAGGCTTTCAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1101
 DB 1021 GTTAAAGGCTTTCAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1080
 QY 1102 TGCAGAGTAAAGCTTTCAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1161
 DB 1081 TGCAGAGTAAAGCTTTCAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1140
 QY 1162 GAGACATATGAT 1221
 DB 1141 GAGACATATGAT 1200
 QY 1222 TCTTTTCAT 1281
 DB 1201 TCTTTTCAT 1260
 QY 1282 ATGAACAAAGTGAAGAGATGAATATATATATATATATATATATATATATATAT 1341
 DB 1261 ATGAACAAAGTGAAGAGATGAATATATATATATATATATATATATATATATAT 1320
 QY 1342 CAATCCACAT 1401
 DB 1321 CAATCCACAT 1380
 QY 1402 TGAAT 1461
 DB 1381 TGAAT 1440
 QY 1462 ATGAT 1515

Db	1441	ANCGATCATATAGAGTGGGTATGAACCTTAAGTAAATGTTTAAATGTGC	15000
Qy	1516	TATATAAGAAAGTGTCTGC	1536
Db	1501	TATATAAAGAAAGTGTCTGC	1521

RESULT 3

AAS84960/c
ID AAS84960 standard; cDNA; 1539 BP.

AC AAS84960;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #20764

KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

XX

100

DR P-PSDB; ABG20773.

PT New isolated poly

PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 1; SEQ ID No 20764; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences, (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques for restoring normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/publ/published.pct_sequences.

Sequence 1539 BP; 396 A; 379 C; 347 G; 417 T; 0 other;

Query Match	83.3%	Score 1280;	DB 23;	Length 1539;
Best Local Similarity	97.3%;	Pred. No. 0;		
Matches 1408; Conservative	0;	Mismatches 23;	Indels 16;	Gaps 10

QY	106	ATCTCTGTGGCCGAGCTCGAGAGCTGGCGCGCCGAGCCCTTGGCCCCCGACGACGAGAGCGG	165
Db	1539	ATGCGCTGTGGCCGAGCTCGAGAGCTGGCGCGCCGAGCCCTTGGCCCCCGACGAGAGCGG	1480
QY	166	GAACGCGAAGCCGAGTCTCCGCGCAACGGGAGCTGCAGATACCTGGGCGAGATCCAAACATCT	225
Db	1479	GAACGCGAAGCCGAGTCTCCGCGCAACGGGAGCTGCAGATACCTGGGCGAGATCCAAACATCT	1420
QY	226	CTCCGCTGCGAGCGTCAAGAAAGACGACCGCACCGGACCCCGGCAACCTGTGCGATTCGCG	285
Db	1419	CTCCGCTGCGAGCGTCAAGAAAGACGACCGCACCGGAGCTGCAGATACCTGGGCGAGTTCGCG	1360
QY	286	ATGCAGGCGCGCTCAAGCTTGAGAGATGAATTCCTCTGTCTGACACAAACGTGTGTC	345
Db	1299	TGGAAGGCTGTATTGGAAGATGTGCTGTGATTATCAAGGAGATCCAAAGCTGAAGAG	1240
QY	406	CTGTCTTCAAGGAGATGAATAATGTGGAGTGCACATGAGATCCCGAGACTTTTGTGACAC	465
Db	1239	CTGTCTTCAAGGAGATGAATAATGTGGAGTGCACATGAGATCCCGAGACTTTTGTGACAC	1180
QY	466	CTGGGATTTCTCCACAGAGAGAGAGGAGACTTGGGCGCACTTATATGCTTCAAGTGAAG	525
Db	1179	CTGGGATTTCTCCACAGAGAGAGAGGAGACTTGGGCGCACTTATATGCTTCAAGTGAAG	1120
QY	526	CATTTTGGGGCAGAAATACAGAGATATGAGATCAATTATTCAGGACAGGAGTGTGACAA	585
Db	1119	CATTTTGGGGCAGAAATACAGAGATATGAGATCAATTATTCAGGACAGGAGTGTGACAA	1060
QY	586	CTGCAAAAGATGATGACACCATCAAAACCAACCTGACACAGAAAGATCATATTTGC	645
Db	1059	CTGCAAAAGATGATGACACCATCAAAACCAACCTGACACAGAAAGATCATATTTGC	1000
QY	646	-GCTTGGATTCAGAGATC-TTCTCTGATGGCGCTGTCTCAATGCCATCCCTGTGCC	703
Db	999	GGCTTGAATTCAGAGATCTTTCTCTGTAGTGGCGCTGTCTCAATGCCATCCCTGTGCC	940
QY	704	AGTTCTATATGATGGAGACGT-GAGCTGTCTGCGCGAGCTGTA-CCAGATATGGGAGACAT	761
Db	939	AGTTCTATATGATGGAGACGTGAGTGTCTCTTCCACACGTACCCGGGATGTGGGAGACAT	880
QY	762	GGGCGCTGGGTGTGCTCTTCCACATGCGCAGCTACGCGCTGTGACGTACATGATGGAGA	821
Db	879	GGGCGCTGGGTGTGCTCTTCCACATGCGCAGCTACGCGCTGTGACGTACATGATGGAGA	820
QY	822	CATCACGGGC-CTGAAGCCAGGTGACTTATACACTTTTGGGAGATGACATATTTAC	880
Db	819	CATCACGGGCTCTAAAGCTAAGTGTCTCTATCCACTCTGGGAGATGACATATTTAC	760
QY	881	TGATATCATATCGAGCCACTGAAATTCAGCTTCAAGG-ATCCGACCTTTCCAA	938
Db	759	TGATATCATATCGAGCCACTGAAATTCAGCTTCAAGGAGAAACCGACCTTTCCAA	700
QY	939	GCTCAGATTCCTCGAAAGTTGAGAAATTTGATGACTTCAAAGCTGAAAGCTTTCAGAT	998
Db	699	GCTCAGATTCCTCGAAAGTTGAGAAATTTGATGACTTCAAAGCTGAAAGCTTTCAGAT	640
QY	999	TGAAGGGTACATCCGACTCCAACTATTAATAATGGAATGCGCTTTAGGGCTTTCCA	1055
Db	639	TGAAGGGTACATCCGACTCCAACTATTAATAATGGAATGCGCTTCTGAAGGCTTTCCA	580
QY	1059	AGGAGCTTGAAGGATTTGTCAGTCTTTAGGGTTGGGCTGAGTGCAGGATAAAGTTC	1118
Db	579	AGGAGCTTGAAGGATTTGTCAGTCTTTAGGGTTGGGCTGAGTGCAGGATAAAGTTC	520
QY	1119	TTTTTGTCTTAAAGAAAGAACTAGCTCAAAATCTGTCCGTGACCTATCAGTTATTT	1178
Db	519	TTTTTGTCTTAAAGAAAGAACTAGCTCAAAATCTGTCCGTGACCTATCAGTTATTT	460

RESULT 4	
ID	ABA93401/c
XX	ABA93401 standard; DNA; 45716 BP.
XX	
AC	ABA93401;
DT	22-APR-2002 (first entry)
XX	
DE	Human RTS-alpha gene SSO ID NO.5.
XX	
KW	Human; RTS-alpha; RTS-beta; RTS; thymidylate synthase; chromosome 18;
KW	mutant thymidylate synthase; enzyme; antineoplastic; antidepressant;
KW	neuroleptic; nootropic; tranquilizer; gene therapy; hypomania;
KW	neuropsychiatric disorder; bipolar affective disorder; schizophrenia;
KW	severe bipolar affective disorder; bipolar affective disorder;
KW	major depression; attention deficit disorder; schizoaffective disorder
XX	gene; ds.
OS	Homo sapiens.
XX	
XX	US6323244-B1.
XX	
PD	27-NOV-2001.
XX	
XX	05-NOV-1997; 97US-0965048.
XX	
PR	05-NOV-1997; 97US-0965048.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
PA	(REGC) UNIV CALIFORNIA.
PI	
PI	Chen H, Freimer NB;
PI	
DR	WPI; 2002-112960/15.
XX	
PT	Treating an RTS (mutant form of thymidylate synthase) - mediated
PT	neuropsychiatric disorder, e.g., bipolar affective disorder, comprises
PT	administering a compound, identified by an RTS assay, that interferes
PT	with RTS gene product -
PS	
PS	Disclosure; Fig 3A-S; 110pp; English.
CC	
CC	The present invention describes a method for treating an RTS (a mutant
CC	form of thymidylate synthase) mediated neuropsychiatric disorder. The

		40.9%;	Score 627.8;	DB 24;	Length 45716;
		Best Local Similarity 91.1%;	Pred. No. 2.9e-153;		
		Matches 665;	Conservative 0;	Mismatches 65;	Indels 0;
					Gaps
QY	807	GTACATGATTCGCAACATCACGGGCTGAAAGCCAGGTGACTTTATATACACTTTGGAGAA	866		
Db	44683	GAACCTTGTGATCACTCCTGTGACTGTTCATGACATGAGGACGATTTACAAAG	44624		
QY	867	TGCATATTTTAACTGAAATCCATCGAGCCACTGAAATTCAGCTTCAGCGGAACCG	926		
Db	44623	GTGTCATATTAAGCAAAATTAAGGCTTATTTTGTGTTTAAAGCTCAGGAGAAACCG	44564		
QY	927	ACCTTCCCAAGGTCAGGATCTTCGAAAAGTTGAGAAAATGTGATCACTTCAAGCTGA	986		
Db	44563	ACCTTCCCAAGGTCAGGATCTTCGAAAAGTTGAGAAAATGTGATCACTTCAAGCTGA	44504		
QY	987	AACATTTCAATTTGAAAGGGTCAATCCGATCCAACTATTAATAATGGAATGCGTGT	1046		
Db	44503	AACATTTCAATTTGAAAGGGTCAATCCGATCCAACTATTAATAATGGAATGCGTGT	44444		
QY	1047	GGGTGCTTTCAAAGGAGCTNGAAGATATTTGTCAAGCTTTAGGGGTTGGGCTGATGCG	1106		
Db	44443	GGGTGCTTTCAAAGGAGCTNGAAGATATTTGTCAAGCTTTAGGGGTTGGGCTGATGCG	44384		
QY	1107	AGGTAAAGTCTCTTTTGCTTAAAGAAANAAGAACTAGTCAAATAATCTGTCGTGAC	1166		
Db	44383	AGGTAAAGTCTCTTTTGCTTAAAGAAANAAGAACTAGTCAAATAATCTGTCGTGAC	44324		
QY	1167	CTATCAGTTATTAATTTTAAAGATGTGTCACATGCGAAATGTAATCTGCGACTTCT	1226		
Db	44323	CTATCAGTTATTAATTTTAAAGATGTGTCACATGCGAAATGTAATCTGCGACTTCT	44264		
QY	1227	CCATATTAAGGCTTTGAGTTAATCACTGAGAGGATCTACAAATGCGAGGTTATGAA	1286		
Db	44263	CCATATTAAGGCTTTGAGTTAATCACTGAGAGGATCTACAAATGCGAGGTTATGAA	44204		
QY	1287	CAAAATGAGGAGATGAATGATATGTCTCTTAAAGAAAATGATATGTGACTTTCAATC	1346		
Db	44203	CAAAATGAGGAGATGAATGATATGTCTCTTAAAGAAAATGATATGTGACTTTCAATC	44144		
QY	1347	CCAGGTACTTAAAGAAGGTGTGTAATTCACAAGCTATTTTGGATATTTTAAAGA	1406		
Db	44143	CCAGGTACTTAAAGAAGGTGTGTAATTCACAAGCTATTTTGGATATTTTAAAGA	44084		
QY	1407	TATTTTAAAGAAATTTGCAAAAGCAATTCCTCCCAATCTGAGGGAGCTGAGTAACCATGA	1466		
Db	44083	TATTTTAAAGAAATTTGCAAAAGCAATTCCTCCCAATCTGAGGGAGCTGAGTAACCATGA	44024		
QY	1467	TCATGATGAGAGTGTGTTTGAATCTTAAAGTTATGTTTATATGTGCTATTAATAAGA	1526		
Db	44023	TCATGATGAGAGTGTGTTTGAATCTTAAAGTTATGTTTATATGTGCTATTAATAAGA	43964		
QY	1527	AGTGTCTGCG	1536		
Db	43963	AGTGTCTGCG	43954		

RESULT 5
 ABA93402/c
 ID ABA93402 standard; DNA; 45989 BP.
 XX ABA93402;
 AC ABA93402;
 XX ABA93402;
 DT 22-APR-2002 (first entry)
 XX 22-APR-2002 (first entry)
 DE Human rTS-beta gene SEQ ID NO:6.
 XX Human rTS-beta gene SEQ ID NO:6.
 KW Human; rTS-alpha; rTS-beta; rTS; thymidylate synthase; chromosome 18;
 KW mutant thymidylate synthase; enzyme; antihuman; antidepressant;
 KW neuroleptic; neurotropic; tranquilizer; gene therapy; hypomania;
 KW neuropsychiatric disorder; bipolar affective disorder; schizophrenia;
 KW severe bipolar affective disorder; bipolar affective disorder;
 KW major depression; attention deficit disorder; schizoaffective disorder;
 KW gene; ds.
 XX Homo sapiens.
 OS Homo sapiens.
 XX US6323244-B1.
 PN US6323244-B1.
 XX 27-NOV-2001.
 PD 27-NOV-2001.
 XX 05-NOV-1997; 97US-0965048.
 PR 05-NOV-1997; 97US-0965048.
 XX 05-NOV-1997; 97US-0965048.
 PR 05-NOV-1997; 97US-0965048.
 XX (MILL-) MILLENNIUM PHARM INC.
 PA (REGC) UNIT CALIFORNIA.
 XX Chen H, Freimer NB;
 PI Chen H, Freimer NB;
 XX WPI; 2002-112960/15.
 DR WPI; 2002-112960/15.
 XX Treating an rTS (mutant form of thymidylate synthase) - mediated
 PT neuropsychiatric disorder, e.g., bipolar affective disorder, comprises
 PT administering a compound, identified by an rTS assay, that interferes
 PT with rTS gene product -
 XX Disclosure; Fig 4A-S; 110pp; English.
 PS The present invention describes a method for treating an rTS (a mutant
 XX form of thymidylate synthase) mediated neuropsychiatric disorder. The
 CC method comprises identifying a compound (C) that can be used to treat
 CC the disorder, by contacting a small molecule test compound (I) with an
 CC rTS protein (II) comprising a sequence of either 361 or 416 amino acids
 CC (AB05587 or AB05588) determining whether (I) binds to (II). (C) has
 CC an anionic, antidepressant, neuroleptic, neurotropic and tranquilizer
 CC activities. rTS polynucleotide sequences can be used in gene therapy.
 CC The method is useful in identifying compounds for treating rTS-mediated
 CC neuropsychiatric disorders including bipolar affective disorder, e.g.,
 CC severe bipolar affective (mood) disorder, bipolar affective (mood)
 CC disorder with hypomania, major depression, schizophrenia, attention
 CC deficit disorder and schizoaffective disorder. The present sequence
 CC represents human rTS-beta gene which is used in the exemplification of
 CC the present invention. rTS has been located to chromosome 18.
 XX Sequence 45989 BP; 12183 A; 10457 C; 10488 G; 12799 T; 62 other;
 SQ Sequence 45989 BP; 12183 A; 10457 C; 10488 G; 12799 T; 62 other;
 Query Match 40.9%; Score 627.8; DB 24; Length 45989;
 Best Local Similarity 91.1%; Pred. No. 2,9e-153;
 Matches 665; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
 QY 807 GTACATGATTGGCAGATCAGCGGCGCTGAAGCCAGGACTTATACACTTTGGGAGA 866
 DB 44956 GAACTTTGATCACTCTGTGACTTTTTCATGACATGAGAGCAATTACACAG 44897
 QY 867 TGACATATTTTACCTGATCAGATGAGCACTGAAATTCAGCTTCAAGGAAACCCAG 926
 DB 44896 GTCTGACATTTATGGGAAATATAGCCTTATTTTGTGTTTACCTTCAAGGAAACCCAG 44837
 QY 927 ACCTTTCCAAAGCTAGGATTTCTTGAAAAGTTGAAAATATGATGACTTCAAAAGCTGA 986

DB 44836 ACCTTTCCAAAGCTAGGATTTCTTGAAAAGTTGAAAATATGATGACTTCAAAAGCTGA 44777
 QY 987 AGACTTTCAAGTTGAAGGGTACATCGCATTCACATTTAAATGGAATAGGCTGTTTA 1046
 DB 44776 AGACTTTCAAGTTGAAGGGTACATCGCATTCACATTTAAATGGAATAGGCTGTTTA 44717
 QY 1047 GGGTGTCTTTCAAAAGGAGCTNAGAGATTTGTCAAGCTTTTAAAGGGTGGGCTGAGATCCG 1106
 DB 44716 GGGTGTCTTTCAAAAGGAGCTNAGAGATTTGTCAAGCTTTTAAAGGGTGGGCTGAGATCCG 44657
 QY 1107 AGGTAAAGTCTTTTGTCTTAAAGAAAGGAATAGGTCAAAATCTGTCCGTGAC 1166
 DB 44656 AGGTAAAGTCTTTTGTCTTAAAGAAAGGAATAGGTCAAAATCTGTCCGTGAC 44597
 QY 1167 CATCAGTTTAAATTTTAAAGATGTTGCCACTGGCAATGTAATGTCAGTTCCTT 1226
 DB 44596 CATCAGTTTAAATTTTAAAGATGTTGCCACTGGCAATGTAATGTCAGTTCCTT 44537
 QY 1227 CCATATTAAGGCTTTGAGTTAATCTCACTGAGGGTATCTGACATGCTGAGTTATGA 1286
 DB 44536 CCATATTAAGGCTTTGAGTTAATCTCACTGAGGGTATCTGACATGCTGAGTTATGA 44477
 QY 1287 CAAAGTGAAGAAATGATATGTCTCTTACCAAAAATGATATGCTATTCATC 1346
 DB 44476 CAAAGTGAAGAAATGATATGTCTCTTACCAAAAATGATATGCTATTCATC 44417
 QY 1347 CACGCTACTTAAAGAGGTTGGATTTCAACAATTTTGGAAATTTTATGA 1406
 DB 44416 CACGCTACTTAAAGAGGTTGGATTTCAACAATTTTGGAAATTTTATGA 44357
 QY 1407 TATTTAAATTTCAAGCTATTCCTCAATCTGAGGAGCTGAGTAACACATGA 1466
 DB 44356 TATTTAAATTTCAAGCTATTCCTCAATCTGAGGAGCTGAGTAACACATGA 44297
 QY 1467 TCATGATGAGAGTGGTTATGAATTTANAGTTGTTTATATGTTGCTATTAATGA 1526
 DB 44296 TCATGATGAGAGTGGTTATGAATTTANAGTTGTTTATATGTTGCTATTAATGA 44237
 QY 1527 AGTGTCTGCG 1536
 DB 44236 AGTGTCTGCG 44227

RESULT 6
 AAF31109
 ID AAF31109 standard; cDNA; 18596 BP.
 XX AAF31109;
 AC AAF31109;
 XX AAF31109;
 DT 27-APR-2001 (first entry)
 DE Thymidylate synthase coding sequence.
 KW Analyte-binding enzyme; analyte analysis; ss.
 XX Homo sapiens.
 OS Homo sapiens.
 PN WO200102600-A2.
 XX WO200102600-A2.
 PD 11-JAN-2001.
 DB 30-JUN-2000; 2000WO-US18057.
 PR 06-JUL-1999; 99US-0347878.
 PR 06-DEC-1999; 99US-0457205.
 PA (GEAT) GEN ATOMICS.
 PI Yuan C;
 XX WPI; 2001-071583/08.

PT Assaying method, useful for prognosis and diagnosis of disease.
PT comprises contacting sample with a mutant analyte-binding enzyme and
detecting binding -

XX Disclosure, Page -, 187pp; English.

XX The present invention relates to a method for assaying an analyte in a
CC sample comprising: contacting the sample with a mutant analyte-binding
CC enzyme which has binding affinity for the analyte or an immediate
CC analyte enzymatic conversion product but has attenuated catalytic
CC activity; and detecting resulting binding. The method is useful in
CC monitoring biological systems/processes, or prognosis/diagnosis of
CC disease caused by imbalances of the analytes. The present sequence is
CC a coding sequence used in the present invention.
CC Note: the present sequence is not shown in the specification, but was
CC from Genbank, using information given in the specification.

XX Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;

Query Match 39.8%; Score 611.8; DB 22; Length 18596;
Best Local Similarity 90.4%; Pred. No. 2.9e-149;
Matches 665; Conservative 0; Mismatches 65; Indels 6; Gaps 1;

QY 807 GTACATGATTGCGCAGATCAAGGCGCTGAAGCCAGTACTTTATACACTTTGGGAGA 866

DB 15510 GAACCTTTGTGATCAGATCCTGTGTACTTTTTCACGAGATGAGGAGATTTCAACAG 15569

QY 867 TGCACATATTTTACCTGATCAATGAGCCACTGAAATTCAGCTTCAGCGAACCAG 926

DB 15570 GTCTGACATATATGCGAAATATATGCGCTTATTTTGTATTACCTTCAGCGAACCAG 15629

QY 927 ACCCTTCCGAAGTCAAGATCTCTCGAAAGTTGAGAAATGTAGACTTCAAGCTGA 986

DB 15630 ACCCTTCCGAAGTCAAGATCTCTCGAAAGTTGAGAAATGTAGACTTCAAGCTGA 15689

QY 987 AGACTTTCAGATTGAAGGGGACATCCCATCCACTATTTAAATGAAATGCGTGTTA 1046

DB 15690 AGACTTTCAGATTGAAGGGGACATCCCATCCACTATTTAAATGAAATGCGTGTTA 15749

QY 1047 GGGTCTTTCAAGAGAGCTNGAAGGATATGTGAGTCTTTAGGGTGTGGATGCGG 1106

DB 15750 GGGTCTTTCAAGAGAGCTNGAAGGATATGTGAGTCTTTAGGGTGTGGATGCGG 15809

QY 1107 AGGTAAAGTCTTTTGTCTTAAAGAAAGAACTAGTCAAAAATCTGTCCGTGAC 1166

DB 15810 AGGTAAAGTCTTTTGTCTTAAAGAAAGAACTAGTCAAAAATCTGTCCGTGAC 15869

QY 1167 CATAGATTATATTTTAAAGATGTTGCCATGCGGAATGTAACTGTGCGAGTCTTT 1226

DB 15870 CATAGATTATATTTTAAAGATGTTGCCATGCGGAATGTAACTGTGCGAGTCTTT 15929

QY 1227 CCATATATTAAGGCTTTAGTTAACTCACTGAGGGTATCTGACATGCTGAGTTATGA 1286

DB 15930 CCATATATTAAGGCTTTAGTTAACTCACTGAGGGTATCTGACATGCTGAGTTATGA 15989

QY 1287 CAAAGTGAAGAAATGAATGTATGTCTTTAGCAAAAACATGTATGTGCAATTCATC 1346

DB 15990 CAAAGTGAAGAAATGAATGTATGTCTTTAGCAAAAACATGTATGTGCAATTCATC 16049

QY 1347 CCAAGTCTTTTAAAGAGTGTGGTAATTCAGAACCTTTTGTGAAATTTTGTGA 1406

DB 16050 CCAAGTCTTTTAAAGAGTGTGGTAATTCAGAACCTTTTGTGAAATTTTGTGA 16109

QY 1407 TATTTTAAGATTTCAACAAGTATCCCTCAAAATCTGAGGAGCTGAGTAACAACACGA 1466

DB 16110 TATTTTAAGATTTCAACAAGTATCCCTCAAAATCTGAGGAGCTGAGTAACAACACGA 16169

QY 1467 TCATGATGTAGAGTGTGGTTATGAAT-----TTAAGTGTGTTTATATGTGCTATA 1520

DB 16170 TCATGATGTAGAGTGTGGTTATGAATTTTAAAGTTATATGTTTATGTGCTATA 16229

QY 1521 TAAAGAGTGTCTGCG 1536

DB 16230 TAAAGAGTGTCTGCG 16245

RESULT 7

XX AAC91215

XX AAC91215 standard; DNA; 18596 BP.

XX AAC91215;

XX 20-MAR-2001 (first entry)

XX Human thymidylate synthase gene SEQ ID NO: 11.

XX Human; schizophrenia; developmental disorder; spina bifida cystica;

XX Tourette's syndrome; bipolar illness; autism; conduct disorder;

XX attention deficit disorder; obsessive compulsive disorder;

XX chronic multiple tic syndrome; learning disorder; polymorphism; ds.

XX Homo sapiens.

XX WO200071754-A1.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000WO-US14354.

XX 25-MAY-1999; 99US-0318448.

XX (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.

XX Johnson WG, Stenroos ES;

XX WPI; 2001-025174/03.

XX Diagnosing a developmental disorder, e.g. schizophrenia, by forming

XX datasets (DS) of genetic (e.g. genotypes of folate metabolism alleles)

XX and environmental variables affecting an individual and then comparing

XX these DS with reference DS -

XX Disclosure, Page 125-131; 156pp; English.

XX The present invention provides a novel method of estimating the

XX susceptibility of an individual to a developmental disorder using genetic

XX and environmental variables. The method can be used in the diagnosis,

XX prevention and treatment of disorders such as schizophrenia, spina bifida

XX cystica, Tourette's syndrome, bipolar illness, autism, conduct disorders,

XX attention deficit disorder, obsessive compulsive disorder, chronic

XX multiple tic syndrome and learning disorders such as dyslexia.

XX Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;

Query Match 39.8%; Score 611.8; DB 22; Length 18596;
Best Local Similarity 90.4%; Pred. No. 2.9e-149;
Matches 665; Conservative 0; Mismatches 65; Indels 6; Gaps 1;

QY 807 GTACATGATTGCGCAGATCAAGGCGCTGAAGCCAGTACTTTATACACTTTGGGAGA 866

DB 15510 GAACCTTTGTGATCAGATCCTGTGTACTTTTTCACGAGATGAGGAGATTTCAACAG 15569

QY 867 TGCACATATTTTACCTGATCAATGAGCCACTGAAATTCAGCTTCAGCGAACCAG 926

DB 15570 GTCTGACATATATGCGAAATATATGCGCTTATTTTGTATTACCTTCAGCGAACCAG 15629

QY 927 ACCCTTCCGAAGTCAAGATCTCTCGAAAGTTGAGAAATGTAGACTTCAAGCTGA 986

DB 15630 ACCCTTCCGAAGTCAAGATCTCTCGAAAGTTGAGAAATGTAGACTTCAAGCTGA 15689

QY 987 AGACTTTCAGATTGAAGGGGACATCCCATCCACTATTTAAATGAAATGCGTGTTA 1046

DB 15690 AGACTTTCAGATTGAAGGGGACATCCCATCCACTATTTAAATGAAATGCGTGTTA 15749

QY 1047 GGGTCTTTCAAGAGAGCTNGAAGGATATGTGAGTCTTTAGGGTGTGGATGCGG 1106

Db 15750 GGGTGGCTTTCAAGAGAGCTGCAAGATATTTGTCACTCTTTAGGGGTTGGGCTGATGCG 15809
 Qy 1107 AGGTAAAGTCTTTTGGCTTAAGAAAGAAAGAACTAGATCAAAATCTGCTCCGTGAC 1166
 Db 15810 AGGTAAAGTCTTTTGGCTTAAGAAAGAAAGAACTAGATCAAAATCTGCTCCGTGAC 15869
 Qy 1167 CTATCAGTTATTAATTTTAAAGATGTGCACTGGCAAAATGTAATCTGCTGCAATCTTT 1226
 Db 15870 CTATCAGTTATTAATTTTAAAGATGTGCACTGGCAAAATGTAATCTGCTGCAATCTTT 15929
 Qy 1227 CCAATTAATAAGGCTTTGAGTTAACTCACTGAGGGTATCTGCAAAATGCTGAGGTTAGAA 1286
 Db 15930 CCAATTAATAAGGCTTTGAGTTAACTCACTGAGGGTATCTGCAAAATGCTGAGGTTAGAA 15989
 Qy 1287 CAAAGTGAAGAGATGAATATATGTGCTTTAGCAAAAATGTAATGCAATTTTGAATC 1346
 Db 15990 CAAAGTGAAGAGATGAATATATGTGCTTTAGCAAAAATGTAATGCAATTTTGAATC 16049
 Qy 1347 CCAAGTGAAGAGATGAATATATGTGCTTTAGCAAAAATGTAATGCAATTTTGAATC 1406
 Db 16050 CCAAGTGAAGAGATGAATATATGTGCTTTAGCAAAAATGTAATGCAATTTTGAATC 16109
 Qy 1407 TATTTAAGAAATTTCAAGCTATTCCTCAAAATCTGAGGAGCTGATTAACCATGGA 1466
 Db 16110 TATTTAAGAAATTTCAAGCTATTCCTCAAAATCTGAGGAGCTGATTAACCATGGA 16169
 Qy 1467 TCATGATGTAGAGTGTGGTTATGAATCTTTTAAAGTGTGTTATATATGTTGCTATTA 1520
 Db 16170 TCATGATGTAGAGTGTGGTTATGAATCTTTTAAAGTGTGTTATATATGTTGCTATTA 16229
 Qy 1521 TAAAGAGTGTCTGC 1536
 Db 16230 TAAAGAGTGTCTGC 16245

RESULT 8
 ABN95092
 ID ABN95092 standard; DNA; 18596 BP.
 XX
 AC ABN95092;
 XX
 XX 13-AUG-2002 (first entry)
 XX
 DE Gene #1590 used to diagnose liver cancer.
 XX
 KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 KW metastatic liver tumor; cytostatic; expression profile; disease state;
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 XX
 OS Homo sapiens.
 XX
 PN WO200229103-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 02-OCT-2001; 2001MO-US30589.
 XX
 XX 02-OCT-2000; 2000US-237054P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 DR WPI; 2002-426119/45.
 XX
 PT Diagnosing and detecting the progression of liver cancer.
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample -
 XX
 PS Claim 1; SEQ ID NO 1590; 298bp; English.
 CC The invention relates to a novel method for diagnosing and detecting the

CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumor in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN95092-ABN9745 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 XX

SQ Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;

Query Match 39.8%; Score 611.8; DB 24; Length 18596;

Best Local Similarity 90.4%; Pred. No. 2,9e-149;

Matches 665; Conservative 0; Mismatches 65; Indels 6; Gaps 1;

Qy 807 GTACATGATTCGCGACATCAAGGCGCTGAGCAAGTGAATTAACACTTTGGGAGA 866
 Db 15510 GAACCTTGTATCATCATCTGTAATCTTTTACAGCAATGAGAGCAATTAACACAG 15569
 Qy 867 TCGACATTAATTAACCTGAATCAATCGAGCACTGAAATTAACCTTCAAGCAAGAACCGAG 926
 Db 15570 GTGCGATTAATTAAGCAAAATTAATGCTTATTTGTTTATGCTTCAAGCAAGAACCGAG 15629
 Qy 927 ACCCTTCCCAAGAGCTGAGATTCCTGAAAAGTGAAGAAAATTAATGATTAACCTTCAAGCTGA 986
 Db 15630 ACCCTTCCCAAGAGCTGAGATTCCTGAAAAGTGAAGAAAATTAATGATTAACCTTCAAGCTGA 15689
 Qy 987 AGACTTTGAGTTGAAGGTTAATTCGATTCGATCAATTAATTAATGAAATGAGCTGTTTA 1046
 Db 15690 AGACTTTGAGTTGAAGGTTAATTCGATTCGATCAATTAATTAATGAAATGAGCTGTTTA 15749
 Qy 1047 GGGTCTTTCAAGAGCTGAGGATTAATGCTGAGCTTTAGGGGTTGGGCTGATGCGG 1106
 Db 15750 GGGTCTTTCAAGAGCTGAGGATTAATGCTGAGCTTTAGGGGTTGGGCTGATGCGG 15809
 Qy 1107 AGGTAAAGTCTTTTGGCTTAAGAAAGAAAGAACTAGATCAAAATCTGCTGATGAC 1166
 Db 15810 AGGTAAAGTCTTTTGGCTTAAGAAAGAAAGAACTAGATCAAAATCTGCTGATGAC 15869
 Qy 1167 CTATCAGTTATTAATTTTAAAGATGTGCACTGGCAAAATGTAATGCAATTTTGAATC 1226
 Db 15870 CTATCAGTTATTAATTTTAAAGATGTGCACTGGCAAAATGTAATGCAATTTTGAATC 15929
 Qy 1227 CCAATTAATAAGGCTTTGAGTTAACTCACTGAGGGTATCTGCAAAATGCTGAGGTTAGAA 1286
 Db 15930 CCAATTAATAAGGCTTTGAGTTAACTCACTGAGGGTATCTGCAAAATGCTGAGGTTAGAA 15989
 Qy 1287 CAAAGTGAAGAGATGAATATATGTGCTTTAGCAAAAATGTAATGCAATTTTGAATC 1346
 Db 15990 CAAAGTGAAGAGATGAATATATGTGCTTTAGCAAAAATGTAATGCAATTTTGAATC 16049
 Qy 1347 CCAAGTGAAGAGATGAATATATGTGCTTTAGCAAAAATGTAATGCAATTTTGAATC 1406
 Db 16050 CCAAGTGAAGAGATGAATATATGTGCTTTAGCAAAAATGTAATGCAATTTTGAATC 16109
 Qy 1407 TATTTAAGAAATTTCAAGCTATTCCTCAAAATCTGAGGAGCTGATTAACCATGGA 1466
 Db 16110 TATTTAAGAAATTTCAAGCTATTCCTCAAAATCTGAGGAGCTGATTAACCATGGA 16169
 Qy 1467 TCATGATGTAGAGTGTGGTTATGAATCTTTTAAAGTGTGTTATATATGTTGCTATTA 1520
 Db 16170 TCATGATGTAGAGTGTGGTTATGAATCTTTTAAAGTGTGTTATATATGTTGCTATTA 16229
 Qy 1521 TAAAGAGTGTCTGC 1536
 Db 16230 TAAAGAGTGTCTGC 16245

Db 181 CCGCCGACGGGAGCTGCACTGAGTACCTGCGGAGGATCCAAACATCTCCGTCGCGGCT 240
QY 241 AGAAGAGACGACCGACCGGACCGGACCGCTGTCGTTATTCGGATGACGAGCCGCTAC 300
Db 241 AGAAGAGACGACCGGACCGGACCGGACCGCTGTCGTTATTCGGATGACGAGCCGCTAC 300
QY 301 AGCTGAGAGATGATTCCTCTGCTGACCAACAGTGTGTCTGGAAGGAGTGTG 360
Db 301 AGCTGAGAGATGATTCCTCTGCTGACCAACAGTGTGTCTGGAAGGAGTGTG 360
QY 361 GAGGAGTGTGCTGTGTTATTCAGGAGATCCAAAGTCTAAAGAGTGTCTTCCAGGGA 420
Db 361 GAGGAGTGTGCTGTGTTATTCAGGAGATCCAAAGTCTAAAGAGTGTCTTCCAGGGA 420
QY 421 GTGGAATCTGGGATGCCAATGATCCGAGACCTTTTGGACAGCTGGGATTCACACC 480
Db 421 GTGGAATCTGGGATGCCAATGATCCGAGACCTTTTGGACAGCTGGGATTCACACC 480
QY 481 AGAAGAGAGGAGCTTGGGACCGATTATGCTTCAGTGAAGCATTTTGGGACGAA 540
Db 481 AGAAGAGAGGAGCTTGGGACCGATTATGCTTCAGTGAAGCATTTTGGGACGAA 540
QY 541 TACAGAGATGAGATCAGATTATTCAGGACAGGAGTGGACACTGCAAGAGTAT 600
Db 541 TACAGAGATGAGATCAGATTATTCAGGACAGGAGTGGACACTGCAAGAGTAT 600
QY 601 GACACCATCAAAACCAACCTTACACAGAGAAATCATCATGTGCGCTTGAATCCAGA 660
Db 601 GACACCATCAAAACCAACCTTACACAGAGAAATCATCATGTGCGCTTGAATCCAGA 660
QY 661 GATCTTCTGATGAGGCGTGCCTCAGACAGCCCTGCGCACTTATGAGTGAAC 720
Db 661 GATCTTCTGATGAGGCGTGCCTCAGACAGCCCTGCGCACTTATGAGTGAAC 720
QY 721 AGTAGAGTCTCTGCGACGCTGACAGAGATCGGAGACATGGGCTCGGTGTGCTTTC 780
Db 721 AGTAGAGTCTCTGCGACGCTGACAGAGATCGGAGACATGGGCTCGGTGTGCTTTC 780
QY 781 AACATCCGACGCTACGCGCTGCTACGTAATGATGCGACATCCGAGCTTGAAGCA 840
Db 781 AACATCCGACGCTACGCGCTGCTACGTAATGATGCGACATCCGAGCTTGAAGCA 840
QY 841 GGTGACTTATACACCTTTGGGAGATGACATTTTACTGAAATGACATCGACCTG 900
Db 841 GGTGACTTATACACCTTTGGGAGATGACATTTTACTGAAATGACATCGACCTG 900
QY 901 AAAATTCAGCTTACGCGAGAACCCGAGACCTTCCCAAGCTCGAGATTCGAAAGT 960
Db 901 AAAATTCAGCTTACGCGAGAACCCGAGACCTTCCCAAGCTCGAGATTCGAAAGT 960
QY 961 GAGAAATGATGATCTTCAAGCTGAAGATTCAGATTGAAGGATCCGATCCA 1020
Db 961 GAGAAATGATGATCTTCAAGCTGAAGATTCAGATTGAAGGATCCGATCCA 1020
QY 1021 ACTATTTAAATGAAATGAGCTGTTTAAAGGCTTTCAAGAGACTGGAAGATTTGCA 1080
Db 1021 ACTATTTAAATGAAATGAGCTGTTTAAAGGCTTTCAAGAGACTGGAAGATTTGCA 1080
QY 1081 GTCTTTAGGGGTTGGGCTGAGATGCCAGAGTAAAGTTCTTTTGTCTTAAAGAAAG 1140
Db 1081 GTCTTTAGGGGTTGGGCTGAGATGCCAGAGTAAAGTTCTTTTGTCTTAAAGAAAG 1140
QY 1141 AACTAGAGTCAAAATCTGTCCGCTGACCTATGATTAATTTTAAAGATGTGCCACT 1200
Db 1141 AACTAGAGTCAAAATCTGTCCGCTGACCTATGATTAATTTTAAAGATGTGCCACT 1200
QY 1201 GCGAAATGTAACTGTGCGAGTCTTTCCATATAAAGGCTTTGAAGTAACTCACTGAG 1260
Db 1201 GCGAAATGTAACTGTGCGAGTCTTTCCATATAAAGGCTTTGAAGTAACTCACTGAG 1260
QY 1261 GTATCTGACATGCTGAGGTTATGAACAAAGTGAAGAGATGAATGAATGATGTGCTT 1320
Db 1261 GTATCTGACATGCTGAGGTTATGAACAAAGTGAAGAGATGAATGAATGATGTGCTT 1320

QY 1321 CAAAAATGATGATGATTCATTCATTCACGTAATTAAGAGGTTGGTATTTGAC 1380
Db 1321 CAAAAATGATGATGATTCATTCATTCACGTAATTAAGAGGTTGGTATTTGAC 1380
QY 1381 AAGCTATTTTGGAAATTTTAAAGATTTTAAAGATTTTCAAGATTTTCCGCAAT 1440
Db 1381 AAGCTATTTTGGAAATTTTAAAGATTTTAAAGATTTTCAAGATTTTCCGCAAT 1440
QY 1441 CTGAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1496
Db 1441 CTGAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1496
RESULT 12
US-60-412-049-350
; Sequence 350, Application US/60412049
; GENERAL INFORMATION:
; APPLICANT: Baker, Joffe B.
; APPLICANT: Cronin, Maureen T.
; APPLICANT: Kiefer, Michael C.
; APPLICANT: Shak, Steve
; APPLICANT: Walker, Michael Graham
; TITLE OF INVENTION: GENE EXPRESSION PROFILING IN BIOPSIED TUMOR TISSUES
; FILE REFERENCE: GENR.004PR
; CURRENT FILING DATE: 2002-09-18
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 350
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-412-049-350
Query Match 90.8%; Score 1394; DB 85; Length 1536;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1494; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGGGGGGGGGGGACCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Db 1 GGGGGGGGGGGGACCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
QY 61 CCGCGCGGACCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db 61 CCGCGCGGACCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 121 TCGAGAGTGGCGCGCGGCTTGGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 180
Db 121 TCGAGAGTGGCGCGCGGCTTGGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 180
QY 181 CCGCGCGAGCGGAGAGTGAATCTGCGGAGATCAATCTCTCGCTGCGCGCTG 240
Db 181 CCGCGCGAGCGGAGAGTGAATCTGCGGAGATCAATCTCTCGCTGCGCGCTG 240
QY 241 AGAAGAGAGAGCGCGAGCGGACCGGACCTGTGCGTATTTGGGATGAGAGCGGCTAC 300
Db 241 AGAAGAGAGAGCGCGAGCGGACCGGACCTGTGCGTATTTGGGATGAGAGCGGCTAC 300
QY 301 AGCTGAGAGATGATTCCTCTGCTGACCAACAGTGTGTCTGGAAGGAGTGTG 360
Db 301 AGCTGAGAGATGATTCCTCTGCTGACCAACAGTGTGTCTGGAAGGAGTGTG 360
QY 361 GAGAGTGTGCTGTGTTATTCAGGAGATCCAAATGCTTAAAGAGTGTCTTCCAGGGA 420
Db 361 GAGAGTGTGCTGTGTTATTCAGGAGATCCAAATGCTTAAAGAGTGTCTTCCAGGGA 420
QY 421 GTGGAATCTGGGATGCCAATGATCCGAGACCTTTTGGACAGCTGGGATTCACACC 480
Db 421 GTGGAATCTGGGATGCCAATGATCCGAGACCTTTTGGACAGCTGGGATTCACACC 480
QY 481 AGAAGAGAGGAGCTTGGGACCGATTATGCTTCAGTGAAGCATTTTGGGACGAA 540

Qy	853	CAACCTTGGAGATGCAATATTTACCTGAATCAACAGCCATCGAAATTCAGCTT	912
Dd	887	CACACTTGGGAGATGCAATATTTACCTGAATCAACAGCCATCGAAATTCAGCTT	946
Qy	913	CAGCAGAACCCGACACTTTCGCCAAGGCTCAGATCTCTGGAAAAGTGAGAAAATGGAT	972
Dd	947	CAGCAGAACCCGACACTTTCGCCAAGGCTCAGATCTCTGGAAAAGTGAGAAAATGGAT	1006
Qy	973	GACCTCAAGGTAAACATTTTCAGATTGAAGGGTCAATCCGATCCAACTATTAATG	1032
Dd	1007	GACCTCAAGGTAAACATTTTCAGATTGAAGGGTCAATCCGATCCAACTATTAATG	1066
Qy	1033	GAAATGCTGTTTAGGGTGCTTTCCAAAGAGCTGAAAGATATTCAGTCTTAGGGGT	1092
Dd	1067	GAAATGCTGTTTAGGGTGCTTTCCAAAGAGCTGAAAGATATTCAGTCTTAGGGGT	1126
Qy	1093	TGGGCTGGATGCCAGGTAAAGTCTTTTGCTTAAGAAAGAACTAGGTCAAA	1152
Dd	1127	TGGGCTGGATGCCAGGTAAAGTCTTTTGCTTAAGAAAGAAAGAACTAGGTCAAA	1186
Qy	1153	AATCTGCCGTGACCTATACAGTTATATTTTAAAGATGATGTCACATCGGCAATGAC	1212
Dd	1187	AATCTGCCGTGACCTATACAGTTATATTTTAAAGATGATGTCACATCGGCAATGAC	1246
Qy	1213	TGGCCAGTTCTTTCCATTAATAAAGGCTTGAGTTACTACTCGAGGATATCGAAT	1272
Dd	1247	TGGCCAGTTCTTTCCATTAATAAAGGCTTGAGTTACTACTCGAGGATATCGAAT	1306
Qy	1273	GCTGAGTTATGAACAAAGTGAGAGATGAATGATGTCCTTAGCAAAAACATGTA	1332
Dd	1307	GCTGAGTTATGAACAAAGTGAGAGATGAATGATGTCCTTAGCAAAAACATGTA	1366
Qy	1333	TGTGATTTCAATCCACGATCTTATTAAGAAGTGTGTAATTTACAGCTATTTTGG	1392
Dd	1367	TGTGATTTCAATCCACGATCTTATTAAGAAGTGTGTAATTTACAGCTATTTTGG	1426
Qy	1393	GAAATTTTGAATATTTTAAAGATTTCACAAGCTATTCCTCAATCTGAGGAGCTG	1452
Dd	1427	GAAATTTTGAATATTTTAAAGATTTCACAAGCTATTCCTCAATCTGAGGAGCTG	1486
Qy	1453	AGTAACACCATCATCATCATGATGATGAGTGTGATGAACTTA	1496
Dd	1487	AGTAACACCATCATCATCATGATGATGAGTGTGATGAACTTA	1530
RESULT 15			
US-09-652-121-6686			
Sequence 6686, Application US/09652121			
GENERAL INFORMATION:			
APPLICANT: Distefano, Peter			
TITLE OF INVENTION: NOVEL NOCLETIC ACID MOLECULES AND USES			
FILE REFERENCE: 1600.1188-001			
CURRENT APPLICATION NUMBER: US/09/652,121			
CURRENT FILING DATE: 2000-08-30			
PRIOR APPLICATION NUMBER: 60/151,123			
PRIOR FILING DATE: 1999-08-30			
NUMBER OF SEQ ID NOS: 7615			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 6686			
LENGTH: 1829			
TYPE: DNA			
ORGANISM: Homo sapiens			
US-09-652-121-6686			

Oy	73	CTGGCCTGACCTCCGATCCCGCCCGCCGCGCGCAGCTGCTGGCCGAGCTCGAGCTCG	132
Db	107	CTTCGACCTGCCTCCGTCCTCCCGCCCGCCGCGCAGTCTGTGGCCGAGCTCG	166
Oy	133	CCGCGGACCTTTGACCCCGCGCGCAGAGAGCGCGAGCCCGATCCGCGCAGCG	192
Db	167	CGCGGACCTTTGACCCCGCGCGCAGAGAGCGCGAGCCCGATCCGCGCAGCG	226
Oy	193	GAGCTGCAGTACTGTGGGAGATCAACACATCTCCGCTCGGAGCTCAGAGAGAGAC	252
Db	227	GAGCTGCAGTACTGTGGGAGATCAACACATCTCCGCTCGGAGCTCAGAGAGAGAC	286
Oy	253	CCGACGCGCACCGCGCACCTCTGTGGATTTGGCAGTACGCGCGCTACGCTGAGAGAT	312
Db	287	CGCAGCGGACACCGGCGACCTCTGTGGATTTGGCAGTACGCGCGCTACGCTGAGAGAT	346
Oy	313	GAAATCCCTCTGCTGACAAACAAACGTGTCTGTGAGAGGGGTGTTTGGAGAGTGTCTG	372
Db	347	GAAATCCCTCTGCTGACAAACAAACGTGTCTGTGAGAGGGGTGTTTGGAGAGTGTCTG	406
Oy	373	TGAGTTTACAAGGAGATCCAAAGCTAAAGAGCTCTTCCAGGAGAGTGAATCTG	432
Db	407	TGAGTTTACAAGGAGATCCAAAGCTAAAGAGCTCTTCCAGGAGAGTGAATCTG	466
Oy	433	GATGCATAGATCCCGAGACTTTTGTGAACAGCTCGGATTTTCCACAGAGAAAGG	492
Db	467	GATGCATAGATCCCGAGACTTTTGTGAACAGCTCGGATTTTCCACAGAGAAAGG	526
Oy	493	GACTTGGGCCCACTTTATGCGCTTCAATGAGAGCATTTTGGGCGAATAACAGAGATG	552
Db	527	GACTTGGGCCCACTTTATGCGCTTCAATGAGAGCATTTTGGGCGAATAACAGAGATG	586
Oy	553	GAAATCAATATTACAGACAGGAGATTGACCAATCGCAAGAGATTTACACATCAAA	612
Db	587	GAAATCAATATTACAGACAGGAGATTGACCAATCGCAAGAGATTTACACATCAAA	646
Oy	613	ACCAACCTTCAGACAGAGAGATATCATGTGCGCTGTGGAATCCAGAGATCTTCTGT	672
Db	647	ACCAACCTTCAGACAGAGAGATATCATGTGCGCTGTGGAATCCAGAGATCTTCTGT	706
Oy	673	ATGGCGTGCCTTCATGCAATGCTTGTGCAGTTCTATGTGGTGAACGTGACTGTCC	732
Db	707	ATGGCGTGCCTTCATGCAATGCTTGTGCAGTTCTATGTGGTGAACGTGACTGTCC	766
Oy	733	TGCGACGCTTACAGAGATCGGAGATGAGGGCTCGGTGTGCTTTCAACATGCGCAG	792
Db	767	TGCGACGCTTACAGAGATCGGAGATGAGGGCTCGGTGTGCTTTCAACATGCGCAG	826
Oy	793	TACGCCCTGCTCACTGATGATGTTGGGACATCAAGGGCTCGAGGCAAGGTACTTATA	852
Db	827	TACGCCCTGCTCACTGATGATGTTGGGACATCAAGGGCTCGAGGCAAGGTACTTATA	886
Oy	853	CAACCTTTGGAGATGCAATATTACTGTGAATCATGAGCCACTGAAAATTCAGCTT	912
Db	887	CAACCTTTGGAGATGCAATATTACTGTGAATCATGAGCCACTGAAAATTCAGCTT	946
Oy	913	CAGCGAGATCCAGACTTTTCCAAAGCTCAGAGATTTCTTGAAAATTGAATAATGAT	972
Db	947	CAGCGAGATCCAGACTTTTCCAAAGCTCAGAGATTTCTTGAAAATTGAATAATGAT	1006
Oy	973	GACTTCAAGGTAAAGCTTTCAATTGAAGGGTACAATCCGATCCAACTATTTAAATG	1032
Db	1007	GACTTCAAGGTAAAGCTTTCAATTGAAGGGTACAATCCGATCCAACTATTTAAATG	1066
Oy	1033	GAAATGCTGTTTGGGTCTTTCAAGAGCTTGAAGATTTGTCAGTCTTTAAGGGT	1092
Db	1067	GAAATGCTGTTTGGGTCTTTCAAGAGCTTGAAGATTTGTCAGTCTTTAAGGGT	1126
Oy	1093	TGGGCTGGAATCCGAGGTAAAGTCTTTTCTCTTAAAGAAAGAAAGACTAGCTCAAA	1150
Db	1127	TGGGCTGGAATCCGAGGTAAAGTCTTTTCTCTTAAAGAAAGAAAGACTAGCTCAAA	1186

QY	1153	AACCTGTCGGTACCACTCACTTATTAATTTTAAAGATGTTGGCACTGGCAAAATGTAC	1212
Db	1187	AACCTGTCGGTACCACTCACTTATTAATTTTAAAGATGTTGGCACTGGCAAAATGTAC	1246
QY	1213	TGTGCGAGTCTCTTCATTAATTAAGAAGCTTGATTACTACAGAGGATCTGACAAAT	1272
Db	1247	TGTGCGAGTCTCTTCATTAATTAAGAAGCTTGATTACTACAGAGGATCTGACAAAT	1306
QY	1273	GGTGAAGTATGAGCAAAAGTGAAGAAATGAATGTATGTGCTTTAGCAAAAACATGA	1332
Db	1307	GGTGAAGTATGAGCAAAAGTGAAGAAATGAATGTATGTGCTTTAGCAAAAACATGA	1366
QY	1333	TGTGCAATTCATCCACGTAACCTTAATTAAGAAGTGTGTGATTTTCAACAAGCTATTTTG	1392
Db	1367	TGTGCAATTCATCCACGTAACCTTAATTAAGAAGTGTGTGATTTTCAACAAGCTATTTTG	1426
QY	1393	GAAATTTTTTAAATATTTTAAAGATTTTCAAGAATTCCTCCGAATCTGAAGGAGCTG	1452
Db	1427	GAAATTTTTTAAATATTTTAAAGATTTTCAAGAATTCCTCCGAATCTGAAGGAGCTG	1486
QY	1453	AGTAAACCATGATCATGATGATGATGAGTGTGTTGAACTTTA	1496
Db	1487	AGTAAACCATGATCATGATGATGATGAGTGTGTTGAACTTTA	1530

Search completed: November 29, 2002, 05:04:42
Job time : 3199.01 secs

QY 443 GATCCGAGACTTTTGGACAGCCTGGATTCTCCACGAGAGAAAGGAGCTTGG393C 502
DB 241 GATCCGAGACTTTTGGACAGCCTGGATTCTCCACGAGAGAGAGGAGCTTGG393C 300
QY 503 CAGTTATGAGCTTCCAGTGGAGGATTTTGGGCGAGAAATACAGATATGAAATCAAGTT 562
DB 301 CAGTTATGAGCTTCCAGTGGAGGATTTTGGGCGAGAAATACAGATATGAAATCAAGTT 360
QY 563 ATTGAGAGAGGAGTTGACCACTGCAAAAGTATGACCACTCAATCAAAACCAAGCTTG 622
DB 361 ATTGAGAGAGGAGTTGACCACTGCAAAAGTATGACCACTCAATCAAAACCAAGCTTG 420
QY 623 ACAGCAGAGAAATCATCATGTCCTTGGATTCAGAGAAATCTTCTCTGATGCGCTG 682
DB 421 ACAGCAGAGAAATCATCATGTCCTTGGATTCAGAGAAATCTTCTCTGATGCGCTG 480
QY 683 CTCGATGCGATGCTTGGCAGCTTCAATGAGTGAACAGTACGCTGCTG 734
DB 481 CTCGATGCGATGCTTGGCAGCTTCAATGAGTGAACAGTACGCTGCTG 532

RESULT 2

US-09-954-531-124
Sequence 124, Application US/09954531
Patent No. US20020165180A1
GENERAL INFORMATION:
APPLICANT: Meaver, Zoe
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
FILE REFERENCE: 689290-77
CURRENT FILING DATE: 2002-05-02
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/233,133
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,009
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,034
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,509
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: US/60/234,567
PRIOR FILING DATE: 2000-09-22
NUMBER OF SEQ ID NOS: 1392
SOFTWARE: PatentIn version 3.0
SEQ ID NO 124
LENGTH: 18596
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-531-124

Query Match 31.7%; Score 487; DB 9; Length 18596;

Best Local Similarity 99.7%; Pred. No. 4.5e-249; Indels 0; Gaps 0;

Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 908 AGCTTCAGGAGAACCCAGACCTTTCCAAAGCTCAGAGATCTTGGAAATTTGAGAAA 967
DB 15611 AGCTTCAGGAGAACCCAGACCTTTCCAAAGCTCAGAGATCTTGGAAATTTGAGAAA 15670
QY 968 TTGATGACTTCAAGCTGAAGCTTTCAGATGAAGGTGACATCCGATCCACTTTTA 1027
DB 15671 TTGATGACTTCAAGCTGAAGCTTTCAGATGAAGGTGACATCCGATCCACTTTTA 15730
QY 1028 AAATGAAATGCTGTTAGGCTGCTTCAAGAGCTGAGATGATATGTCAGTCTTTA 1087
DB 15731 AAATGAAATGCTGTTAGGCTGCTTCAAGAGCTGAGATGATATGTCAGTCTTTA 15790
QY 1088 GGGGTTGGGCTGATGCCAGAGTAAAGTTCTTTTCTTAAAGAAAGAAAGAACTAGG 1147
DB 15791 GGGGTTGGGCTGATGCCAGAGTAAAGTTCTTTTCTTAAAGAAAGAAAGAACTAGG 15850
QY 1148 TCAGAAATCTGTCGATGACATGATTAATTTTAAAGATGTGCTGCACTGCAAT 1207

DB 15851 TCAGAAATCTGTCGATGACATGATTAATTTTAAAGATGTGCTGCACTGCAAT 15910
QY 1208 GTAAGTGGCAATCTTTCCATTAATAAAGCTTTGATTAATCACTGAGGATCTG 1267
DB 15911 GTAAGTGGCAATCTTTCCATTAATAAAGCTTTGATTAATCACTGAGGATCTG 15970
QY 1268 ACAATGCTGAGTTATGAACAAGTGAAGAAATGAATGATGCTCTTGAAGAAAAC 1327
DB 15971 ACAATGCTGAGTTATGAACAAGTGAAGAAATGAATGATGCTCTTGAAGAAAAC 16030
QY 1328 ATGTATGCAATTCATCCAGTACTTATTAAGAGGTTGATTAATCAAGCTAT 1387
DB 16031 ATGTATGCAATTCATCCAGTACTTATTAAGAGGTTGATTAATCAAGCTAT 16090
QY 1388 TTTTGAATATTTTAAATATTTTAAATTTTCAAGCTATTCCTCAATCTGAGG 1447
DB 16091 TTTTGAATATTTTAAATATTTTAAATTTTCAAGCTATTCCTCAATCTGAGG 16150
QY 1448 AGCTGAGTAAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1496
DB 16151 AGCTGAGTAAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 16199

RESULT 3

US-09-954-531-348
Sequence 348, Application US/09954531
Patent No. US20020165180A1
GENERAL INFORMATION:
APPLICANT: Meaver, Zoe
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using
FILE REFERENCE: 689290-77
CURRENT FILING DATE: 2002-05-02
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/233,133
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,009
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,034
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,509
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: US/60/234,567
PRIOR FILING DATE: 2000-09-22
NUMBER OF SEQ ID NOS: 1392
SOFTWARE: PatentIn version 3.0
SEQ ID NO 348
LENGTH: 18596
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-531-348

Query Match 31.7%; Score 487; DB 9; Length 18596;

Best Local Similarity 99.7%; Pred. No. 4.5e-249; Indels 0; Gaps 0;

Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 908 AGCTTCAGGAGAACCCAGACCTTTCCAAAGCTCAGAGATCTTGGAAATTTGAGAAA 967
DB 15611 AGCTTCAGGAGAACCCAGACCTTTCCAAAGCTCAGAGATCTTGGAAATTTGAGAAA 15670
QY 968 TTGATGACTTCAAGCTGAAGCTTTCAGATGAAGGTGACATCCGATCCACTTTTA 1027
DB 15671 TTGATGACTTCAAGCTGAAGCTTTCAGATGAAGGTGACATCCGATCCACTTTTA 15730
QY 1028 AAATGAAATGCTGTTAGGCTGCTTCAAGAGCTGAGATGATATGTCAGTCTTTA 1087
DB 15731 AAATGAAATGCTGTTAGGCTGCTTCAAGAGCTGAGATGATATGTCAGTCTTTA 15790
QY 1088 GGGGTTGGGCTGATGCCAGAGTAAAGTTCTTTTCTTAAAGAAAGAAAGAACTAGG 1147
DB 15791 GGGGTTGGGCTGATGCCAGAGTAAAGTTCTTTTCTTAAAGAAAGAAAGAACTAGG 15850
QY 1148 TCAGAAATCTGTCGATGACATGATTAATTTTAAAGATGTGCTGCACTGCAAT 1207

	Qy	908	ACCTTCAGCGAGAACCCAGACCTTTCCCAAAGCTACGAGTCTCTTGAAAAGTTGAGAAA	967
Qy	15611	AGCTTCAGCGAGAACCCAGACCTTTCCCAAAGCTACGAGTCTCTTGAAAAGTTGAGAAA	15670	
Dy	966	TTGATGACCTTCAAACTGAGTGAACCTTCAGATTGAAAGGGTACAAATCCGATCCAACTAATA	1027	
Qy	15671	TTGATGACCTTCAAAAGTGAAACCTTCAGATTGAAAGGGTACAAATCCGATCCAACTAATA	15730	
Dy	1028	AAATGGAATGAGCTGTTTAGGGTCCTTCAAGGAGCTGAAAGATATATGTCACTCTTAA	1087	
Qy	15731	AAATGGAATGAGCTGTTTAGGGTCCTTCAAGGAGCTGAAAGATATATGTCACTCTTAA	15790	
Dy	1088	GAGGTTGAGCTGAGATGCGAGAGTAAAGTTCTTTTCTCTTAAAGAAAGAGACTAGG	1147	
Qy	15791	GAGGTTGAGCTGAGATGCGAGAGTAAAGTTCTTTTCTCTTAAAGAAAGAGACTAGG	15850	
Dy	1148	TCAAAATATCTGTCGGAGCACTACAGTTTATTTTAAAGATGTGCCACTGGCAAT	1207	

[illegible]


```
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PCR primer
US-09-797-163-2

Query Match          1.4%; Score 22; DB 10; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.093;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 781 AACATGCCAGCTAGCCCTGC 802
DB 1 AACATGCCAGCTAGCCCTGC 22

RESULT 9
US-09-998-333A-9
; Sequence 9, Application US/09998333A
; Patent No. US20020142328A1
; GENERAL INFORMATION:
; APPLICANT: Kathleen D. Danenberg
; TITLE OF INVENTION: METHOD OF DETERMINING A CHEMOTHERAPEUTIC
; TITLE OF INVENTION: REGIMEN BY ASSAYING GENE EXPRESSION IN PRIMARY TUMORS
; FILE REFERENCE: 11220/146
; CURRENT APPLICATION NUMBER: US/09/998,333A
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: To Be Assigned
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/879,217
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 09/877,178
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/877,177
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 60/250,472
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 60/250,120
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide Primer
US-09-998-333A-9

Query Match          1.4%; Score 22; DB 10; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.093;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 781 AACATGCCAGCTAGCCCTGC 802
DB 1 AACATGCCAGCTAGCCCTGC 22

RESULT 10
US-09-797-163-3/c
; Sequence 3, Application US/09797163
; Patent No. US2002009795A1
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; APPLICANT: Danenberg, Kathleen D.
; APPLICANT: Swenson, Steven
; TITLE OF INVENTION: ISOLATION OF RNA, DNA AND PROTEINS FROM
; TITLE OF INVENTION: FORMALIN-FIXED PARAFFIN-EMBEDDED TISSUE SPECIMENS
; FILE REFERENCE: 13761-770
```

```
; CURRENT APPLICATION NUMBER: US/09/797,163
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: US 09/469,338
; PRIOR FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PCR primer
US-09-797-163-3

Query Match          1.4%; Score 21; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 805 ACGTACATGATTGGCGCATC 825
DB 21 ACGTACATGATTGGCGCATC 1

RESULT 11
US-09-998-333A-8/c
; Sequence 8, Application US/09998333A
; Patent No. US20020142328A1
; GENERAL INFORMATION:
; APPLICANT: Kathleen D. Danenberg
; TITLE OF INVENTION: METHOD OF DETERMINING A CHEMOTHERAPEUTIC
; TITLE OF INVENTION: REGIMEN BY ASSAYING GENE EXPRESSION IN PRIMARY TUMORS
; FILE REFERENCE: 11220/146
; CURRENT APPLICATION NUMBER: US/09/998,333A
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: To Be Assigned
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/879,217
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 09/877,178
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/877,177
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 60/250,472
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 60/250,120
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide Primer
US-09-998-333A-8

Query Match          1.4%; Score 21; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 805 ACGTACATGATTGGCGCATC 825
DB 21 ACGTACATGATTGGCGCATC 1

RESULT 12
US-09-822-485-34
; Sequence 34, Application US/09822485
; Patent No. US20020001825A1
; GENERAL INFORMATION:
; APPLICANT: Itch, No. US20020001825A1uyuki
; TITLE OF INVENTION: No. US20020001825A1el Fibroblast Growth Factor-Like Polypeptides
; FILE REFERENCE: 08035.0001-01000
```

;; CURRENT APPLICATION NUMBER: US/09/822,485
;; CURRENT FILING DATE: 2001-04-02
;; NUMBER OF SEQ ID NOS: 35
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 34
;; LENGTH: 995
;; TYPE: DNA
;; ORGANISM: Mus sp.
US-09-822-485-34

Query Match 1.3%; Score 20; DB 10; Length 995;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGACCACTTG 20
DB 39 GGGGGGGGGGGGACCACTTG 58

RESULT 13

US-09-954-197-1
; Sequence 1, Application US/09954197
; Patent No. US20020107379A1
; GENERAL INFORMATION:
; APPLICANT: MAX, Achim
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR THE CHYA GENE
; FILE REFERENCE: 032301 WD 201
; CURRENT APPLICATION NUMBER: US/09/954,197
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentln version 3.1
; SEQ ID NO 1
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (231)..(1028)
; OTHER INFORMATION:
US-09-954-197-1

Query Match 1.3%; Score 20; DB 10; Length 1200;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 769 GGTGTGCTTTCACATCGC 788
DB 756 GGTGTGCTTTCACATCGC 775

RESULT 14

US-09-987-446-1
; Sequence 1, Application US/09987446
; Patent No. US20020115162A1
; GENERAL INFORMATION:
; APPLICANT: FARWICK, Mike, et al.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR THE CYSQ GENE
; FILE REFERENCE: 032301 WD 245
; CURRENT APPLICATION NUMBER: US/09/987,446
; CURRENT FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentln version 3.1
; SEQ ID NO 1
; LENGTH: 2730
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1014)..(1769)
; OTHER INFORMATION:
US-09-987-446-1

Query Match 1.3%; Score 20; DB 10; Length 2730;

Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 769 GGTGTGCTTTCACATCGC 788
DB 2316 GGTGTGCTTTCACATCGC 2335

RESULT 15

US-09-815-343-185/C
; Sequence 185, Application US/09815343
; Patent No. US2001005596A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine
; APPLICANT: Xu, Jiangchun
; APPLICANT: King, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.504
; CURRENT APPLICATION NUMBER: US/09/815,343
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 185
; LENGTH: 358
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(358)
; OTHER INFORMATION: n = A,T,C or G
US-09-815-343-185

Query Match 1.2%; Score 19; DB 10; Length 358;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1430 TTCCCTGAATCTGAGGA 1448
DB 297 TTCCCTGAATCTGAGGA 279

Search completed: November 29, 2002, 05:48:25
Job time : 631.315 secs

QY 121 TCGAGCTGCGCGCGCGCGCTTTCGCCCCCGCGCACAAGAGCGGAGACGCGCGCGCT 180
 DB 121 TCGAGCTGCGCGCGCGCGCGCTTTCGCCCCCGCGCACAAGAGCGGAGACGCGCGCGCT 180
 QY 181 CGCGCGAGCGGAGAGCTGCACTGACCTGCGGAGAGTCCACACATCTTCGCGTGCAGCGCTC 240
 DB 181 CGCGCGAGCGGAGAGCTGCACTGACCTGCGGAGAGTCCACACATCTTCGCGTGCAGCGCTC 240
 QY 241 AGAAGAGAGCAGCCGACAGCGGACCGGACCCCTGCGGATTCGCGATGACGCGCGCTAC 300
 DB 241 AGAAGAGAGCAGCCGACAGCGGACCGGACCCCTGCGGATTCGCGATGACGCGCGCTAC 300
 QY 301 AGCCTGAGAGATGAATTCCTCTGCTGACACCAAACTGTGTCTGGAAGGCTGTTTG 360
 DB 301 AGCCTGAGAGATGAATTCCTCTGCTGACACCAAACTGTGTCTGGAAGGCTGTTTG 360
 QY 361 GAGGAGTGTGCTGCTGCTTATCAAGGAGTCCAAAGGCTAAAGGCTGCTTCCAGAGGA 420
 DB 361 GAGGAGTGTGCTGCTGCTTATCAAGGAGTCCAAAGGCTAAAGGCTGCTTCCAGAGGA 420
 QY 421 GTGAAATCTGGAGTGCCTAATGATCCGAGACCTTTTGGACAGCGCTGGAATCTCCACC 480
 DB 421 GTGAAATCTGGAGTGCCTAATGATCCGAGACCTTTTGGACAGCGCTGGAATCTCCACC 480
 QY 481 AGAAGAGAGCGGAGCTTGGGCGCACTTTATGCTTCCAGTGGAGCAATTTTGGGCGAGAA 540
 DB 481 AGAAGAGAGCGGAGCTTGGGCGCACTTTATGCTTCCAGTGGAGCAATTTTGGGCGAGAA 540
 QY 541 TACGAGATATGGAATCAGATTTATTCAGGACAGGAGGTTGACCACTGCAAAAGAGATT 600
 DB 541 TACGAGATATGGAATCAGATTTATTCAGGACAGGAGGTTGACCACTGCAAAAGAGATT 600
 QY 601 GACACCATCAAAACCAACCGCTGACGACAGAGAGATCATGAGTGGCTGGAATCCAGAA 660
 DB 601 GACACCATCAAAACCAACCGCTGACGACAGAGAGATCATGAGTGGCTGGAATCCAGAA 660
 QY 661 GATCTCTCTGATGCGGCTGCTGCTTCAATGCGATGCCCTGCGAGGTTCTAATGCTGAAC 720
 DB 661 GATCTCTCTGATGCGGCTGCTGCTTCAATGCGATGCCCTGCGAGGTTCTAATGCTGAAC 720
 QY 721 AGTGAAGTCTGCTGCGAGCTGTACCAAGAGTCCGAGAGATCGGAGATCGGCTGCTTTC 780
 DB 721 AGTGAAGTCTGCTGCGAGCTGTACCAAGAGTCCGAGAGATCGGAGATCGGCTGCTTTC 780
 QY 781 AACCTGCGAGCTAGGCGCTGCTACGTAATGATGAGGAGATCATGAGGCTGGAAGCA 840
 DB 781 AACCTGCGAGCTAGGCGCTGCTACGTAATGATGAGGAGATCATGAGGCTGGAAGCA 840
 QY 841 GGTGACTTTATACACCTTTGGAGATGACATATTTACCTGATCATCATGAGCGACTG 900
 DB 841 GGTGACTTTATACACCTTTGGAGATGACATATTTACCTGATCATCATGAGCGACTG 900
 QY 901 AAAATTCAGCTTCAAGGAGAACCCAGACCTTCCAAAGCTCAGATTTCTCCAAAGTT 960
 DB 901 AAAATTCAGCTTCAAGGAGAACCCAGACCTTCCAAAGCTCAGATTTCTCCAAAGTT 960
 QY 961 GAGAAATGATGATGATCAAGCTGAGAGACTTCAATGAGAGGATCAATCCGACTCA 1020
 DB 961 GAGAAATGATGATGATCAAGCTGAGAGACTTCAATGAGAGGATCAATCCGACTCA 1020
 QY 1021 ACTATTAATGAAATGAGATGCTTTAGGCTGCTTCAAGAGCTGGAAGATTTGTCA 1080
 DB 1021 ACTATTAATGAAATGAGATGCTTTAGGCTGCTTCAAGAGCTGGAAGATTTGTCA 1080
 QY 1081 GTCTTAAAGGAGTGGGCTGAGATGCGAGGTAAGATCTTTTGTCTTAAAGAAAGAG 1140
 DB 1081 GTCTTAAAGGAGTGGGCTGAGATGCGAGGTAAGATCTTTTGTCTTAAAGAAAGAG 1140
 QY 1141 AACTAGGTCAAAATCTGTCGCTGACCTATCAGTTATTTAATTTAAGAGATTTGCACT 1200
 DB 1141 AACTAGGTCAAAATCTGTCGCTGACCTATCAGTTATTTAATTTAAGAGATTTGCACT 1200
 QY 1201 GGCAAATGTAATCTGTCAGCTTCTTCCATTAATTAAGAGCTTTGATTAACCTCACTGAG 1260

DB 1201 GGCAAATGTAATCTGTCAGCTTCTTCCATTAATTAAGAGCTTTGATTAACCTCACTGAG 1260
 QY 1261 GTATCTGACAAATGCTGAGTTATGAGAAAGTGAAGAGATGAATGATGCTTAG 1320
 DB 1261 GTATCTGACAAATGCTGAGTTATGAGAAAGTGAAGAGATGAATGATGCTTAG 1320
 QY 1321 CAAAAATGATGATGATTCATTCACAGTACTTATTAAGAGAGGTTGGAATTTAC 1380
 DB 1321 CAAAAATGATGATGATTCATTCACAGTACTTATTAAGAGAGGTTGGAATTTAC 1380
 QY 1381 AAGCTATTTTGAATTTTGAATATTTAAGATTTTCAAGCTTATTCCTCAAT 1440
 DB 1381 AAGCTATTTTGAATTTTGAATATTTAAGATTTTCAAGCTTATTCCTCAAT 1440
 QY 1441 CTGAGGAGAGCTGATTAACCATCATCATGATGATGAGAGTGTGATTAATTTA 1496
 DB 1441 CTGAGGAGAGCTGATTAACCATCATCATGATGATGAGAGTGTGATTAATTTA 1496

RESULT 3
 US-09-347-878-29
 ; Sequence 29, Application US/09347878C
 ; Patent No. 6376210
 ; GENERAL INFORMATION:
 ; APPLICANT: Yuan, Chong
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
 ; FILE REFERENCE: 25885-1651
 ; CURRENT APPLICATION NUMBER: US/09/347,878C
 ; NUMBER OF SEQ ID NOS: 75
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 29
 ; LENGTH: 942
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Human thymidylate synthase gene: exons 1-8
 ; DATABASE ACCESSION NUMBER: D00596/GenBank
 US-09-347-878-29

Query Match 61.3%; Score 942; DB 4; Length 942;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 942; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 106 ATGCTTGTGGGCTTGGAGAGCTGCGCGCGCGCTTTCGCCCCCGCGCACAAGAGCGG 165
 DB 1 ATGCTTGTGGGCTTGGAGAGCTGCGCGCGCGCGCTTTCGCCCCCGCGCACAAGAGCGG 60
 QY 166 GACGCCGAGCGGCTGCGCGCGGACGTCAGATCCTGAGGAGATCAACATC 225
 DB 61 GACGCCGAGCGGCTGCGCGCGGAGAGTGAATCCTGAGGAGATCAACATC 120
 QY 226 CTCGCTGCGGCTCAGAGAGCAGCCGACAGCGGACCGGACCTGTGCGATTCGCG 285
 DB 121 CTCGCTGCGGCTCAGAGAGCAGCCGACAGCGGACCGGACCTGTGCGATTCGCG 180
 QY 286 ATGCAAGGCGGCTACAGCTGAGAGATGATTCCTCTGCTGACACCAACGTTGTT 345
 DB 181 ATGCAAGGCGGCTACAGCTGAGAGATGATTCCTCTGCTGACACCAACGTTGTT 240
 QY 346 TCGAAGGCTGTTTGGAGAGGCTGCTGCTGCTTATCAAGGATCCAAATGCTAAAGAG 405
 DB 241 TCGAAGGCTGTTTGGAGAGGCTGCTGCTGCTTATCAAGGATCCAAATGCTAAAGAG 300
 QY 406 CTGTCTTCAAGGAGATGAATCTGAGATGCCATGATCCGAGACTTTTGAACAGC 465
 DB 301 CTGTCTTCAAGGAGATGAATCTGAGATGCCATGATCCGAGACTTTTGAACAGC 360
 QY 466 CTGGAATTCCTCCAGAGAGAGAGGAGGACTTGGGCGCAGTTATGCTTCCAGTGAAG 525
 DB 361 CTGGAATTCCTCCAGAGAGAGAGGAGGACTTGGGCGCAGTTATGCTTCCAGTGAAG 420

QY 1208 GTAACGTGCGAAGTCTTCCATATATAAAGCTTGGACTTACTACTGAGGATATCTG 1267
 Db 44282 GTAACGTGCGAAGTCTTCCATATATAAAGCTTGGACTTACTACTGAGGATATCTG 44223
 QY 1268 ACAATGCTGAGGTTATGAACAAGAGGAGAGATGAATGAATGTATGCTTGAAGCAAAAC 1327
 Db 44222 ACAATGCTGAGGTTATGAACAAGAGGAGAGATGAATGAATGTATGCTTGAAGCAAAAC 44163
 QY 1328 ATGATGCTGATTTCAATCCACGACTTATATAAAGAGGTGGATTTCAACAAGCTAT 1387
 Db 44162 ATGATGCTGATTTCAATCCACGACTTATATAAAGAGGTGGATTTCAACAAGCTAT 44103
 QY 1388 TTTTGGAAATTTTGAATATTTTGAATTTTGAATTTTCAAGCTATTTCCCTCAATCTGAGG 1447
 Db 44102 TTTTGGAAATTTTGAATATTTTGAATTTTGAATTTTCAAGCTATTTCCCTCAATCTGAGG 44043
 QY 1448 AGCTGAGTAACACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1436
 Db 44042 AGCTGAGTAACACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 43994

RESULT 6

US-08-965-048-6/c
 ; Sequence 6, Application US/08965048
 ; Patent No. 6323244
 ; GENERAL INFORMATION:
 ; APPLICANT: Chen, Hong
 ; APPLICANT: Freeman, Nelson
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
 ; TITILE OF INVENTION: TREATMENT OF NEUROPSYCHIATRIC DISORDERS
 ; FILE REFERENCE: 7853-093
 ; CURRENT APPLICATION NUMBER: US/08/965,048
 ; CURRENT FILING DATE: 1997-11-05
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 6
 ; LENGTH: 45989
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-08-965-048-6

Query Match 31.7%; Score 487; DB 4; Length 45989;
 Best Local Similarity 99.7%; Pred. No. 4,1e-24;
 Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 908 AGCTTGACGAGAACCCAGACCTTTCCCAAGCTCGAGATTTTGGAAAAGTTGAGAAA 967
 Db 44855 AGCTTGACGAGAACCCAGACCTTTCCCAAGCTCGAGATTTTGGAAAAGTTGAGAAA 44796
 QY 968 TTGATGACTTCAAGCTGAAGCTTTCAGATGAGGATGAGGATCAATCCGATCCAACTATTA 1027
 Db 44795 TTGATGACTTCAAGCTGAAGCTTTCAGATGAGGATGAGGATCAATCCGATCCAACTATTA 44736
 QY 1028 AAATGGAATGCTGTTTAAAGGCTTCAAGAGCTTGAAGATATGTCAGTCTTTA 1087
 Db 44735 AAATGGAATGCTGTTTAAAGGCTTCAAGAGCTTGAAGATATGTCAGTCTTTA 44676
 QY 1088 GGGGTGGGCTGAGTCCGAGTAAAGCTTTTTCCTTAAAGATTAAGAAAGAACTAGG 1147
 Db 44675 GGGGTGGGCTGAGTCCGAGTAAAGCTTTTTCCTTAAAGATTAAGAAAGAAAGAACTAGG 44616
 QY 1148 TCAAAAATCTGTCGAGCTTTCATATATAAAGCTTTGAAGTATGCTGAGGAGTATCTG 1207
 Db 44615 TCAAAAATCTGTCGAGCTTTCATATATAAAGCTTTGAAGTATGCTGAGGAGTATCTG 44556
 QY 1208 GTAACGTGCGAAGTCTTCCATATATAAAGCTTGGACTTACTACTGAGGATATCTG 1267
 Db 44555 GTAACGTGCGAAGTCTTCCATATATAAAGCTTGGACTTACTACTGAGGATATCTG 44496
 QY 1268 ACAATGCTGAGGTTATGAACAAGAGGAGAGATGAATGAATGTATGCTTGAAGCAAAAC 1327
 Db 44495 ACAATGCTGAGGTTATGAACAAGAGGAGAGATGAATGAATGTATGCTTGAAGCAAAAC 44436

QY 1328 ATGATGCTGATTTCAATCCACGACTTATATAAAGAGGTGGATTTCAACAAGCTAT 1387
 Db 44435 ATGATGCTGATTTCAATCCACGACTTATATAAAGAGGTGGATTTCAACAAGCTAT 44376
 QY 1388 TTTTGGAAATTTTGAATATTTTGAATTTTGAATTTTCAAGCTATTTCCCTCAATCTGAGG 1447
 Db 44375 TTTTGGAAATTTTGAATATTTTGAATTTTGAATTTTCAAGCTATTTCCCTCAATCTGAGG 44316
 QY 1448 AGCTGAGTAACACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1436
 Db 44315 AGCTGAGTAACACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 44267

RESULT 7

US-08-965-048-1/c
 ; Sequence 1, Application US/08965048
 ; Patent No. 6323244
 ; GENERAL INFORMATION:
 ; APPLICANT: Chen, Hong
 ; APPLICANT: Freeman, Nelson
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
 ; TITILE OF INVENTION: TREATMENT OF NEUROPSYCHIATRIC DISORDERS
 ; FILE REFERENCE: 7853-093
 ; CURRENT APPLICATION NUMBER: US/08/965,048
 ; CURRENT FILING DATE: 1997-11-05
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1817
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (67)..(1149)
 ; US-08-965-048-1

Query Match 10.9%; Score 168; DB 4; Length 1817;
 Best Local Similarity 99.5%; Pred. No. 8,6e-77;
 Matches 218; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 908 AGCTTGACGAGAACCCAGACCTTTCCCAAGCTCGAGATTTTGGAAAAGTTGAGAAA 967
 Db 1515 AGCTTGACGAGAACCCAGACCTTTCCCAAGCTCGAGATTTTGGAAAAGTTGAGAAA 1456
 QY 968 TTGATGACTTCAAGCTGAAGCTTTCAGATGAGGATGAGGATCAATCCGATCCAACTATTA 1027
 Db 1455 TTGATGACTTCAAGCTGAAGCTTTCAGATGAGGATGAGGATCAATCCGATCCAACTATTA 1396
 QY 1028 AAATGGAATGCTGTTTAAAGGCTTTCAGAGCTTGAAGATATGTCAGTCTTTA 1087
 Db 1395 AAATGGAATGCTGTTTAAAGGCTTTCAGAGCTTGAAGATATGTCAGTCTTTA 1336
 QY 1088 GGGGTGGGCTGAGTCCGAGTAAAGCTTTTTCCTTAAAGATTAAGAAAGAACTAGG 1126
 Db 1335 GGGGTGGGCTGAGTCCGAGTAAAGCTTTTTCCTTAAAGATTAAGAAAGAACTAGG 1297

RESULT 8

US-09-367-007C-7
 ; Sequence 7, Application US/09367007C
 ; Patent No. 6416987
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertino, Joseph R.
 ; APPLICANT: Banerjee, Debabrata
 ; APPLICANT: Tom, Youzhi
 ; APPLICANT: Liu, Chen, Xinyue
 ; TITLE OF INVENTION: Mutants of Thymidylate Synthase and Uses Thereof
 ; FILE REFERENCE: D5978
 ; CURRENT APPLICATION NUMBER: US/09/367,007C
 ; CURRENT FILING DATE: 1999-10-15
 ; PRIOR APPLICATION NUMBER: PCT/US98/02145
 ; PRIOR FILING DATE: 1998-01-03

NUMBER OF SEQ ID NOS: 39
SEQ ID NO 7
LENGTH: 26
TYPE: DNA
ORGANISM: artificial sequence
FEATURE:
NAME/KEY: primer bind
LOCATION: 594..616
OTHER INFORMATION: Sense primer hts-3a for PCR amplification of part
OTHER INFORMATION: of 890 base pair fragment of human TS gene
OTHER INFORMATION: from nucleotide 50 to the C-terminus
US-09-367-007C-7

Query Match 1.6%; Score 26; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 699 CTGCCAGTCTATGTGTGGAACAGTG 724
Db 1 CTGCCAGTCTATGTGTGGAACAGTG 26

RESULT 9
US-09-089-195-26/C
Sequence 26, Application US/09089195
Patent No. 6087489
GENERAL INFORMATION:
APPLICANT: Dean, Nicholas M.
TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION
TITLE OF INVENTION: OF HUMAN THYMIDYLATE SYNTHASE EXPRESSION
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Jane Massey Licata
STREET: 66 East Main Street
CITY: Marlton
STATE: NJ
COUNTRY: USA
ZIP: 08053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/089,195
FILING DATE: herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: ISPH-0309
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 810-1515
TELEFAX: (609) 810-1454
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-09-089-195-26

Query Match 1.6%; Score 25; DB 3; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1036 ATGGCTGTTAGGAGTCTTCAAG 1060
Db 25 ATGGCTGTTAGGAGTCTTCAAG 1

RESULT 10
US-09-367-007C-11

Sequence 11, Application US/09367007C
Patent No. 6416387
GENERAL INFORMATION:
APPLICANT: Bertino, Joseph R.
APPLICANT: Banerjee, Debabrata
APPLICANT: Tong, Youzhi
APPLICANT: Liu-Chen, Xinyue
TITLE OF INVENTION: Mutants of Thymidylate Synthase and Uses Thereof
FILE REFERENCE: D5978
CURRENT APPLICATION NUMBER: US/09/367,007C
CURRENT FILING DATE: 1999-10-15
PRIORITY FILING DATE: PCT/US98/02145
PRIORITY FILING DATE: 1998-01-03
NUMBER OF SEQ ID NOS: 39
SEQ ID NO 11
LENGTH: 25
TYPE: DNA
ORGANISM: artificial sequence
FEATURE:
NAME/KEY: primer bind
LOCATION: 451..475
OTHER INFORMATION: Sense primer hts-5a for PCR amplification of part
OTHER INFORMATION: of 890 base pair fragment of human TS gene
OTHER INFORMATION: from nucleotide 50 to the C-terminus
US-09-367-007C-11

Query Match 1.6%; Score 25; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 556 TCAGATTATTCAGACAGGAGGTTG 580
Db 1 TCAGATTATTCAGACAGGAGGTTG 25

RESULT 11
US-09-367-007C-6/C
Sequence 6, Application US/09367007C
Patent No. 6416987
GENERAL INFORMATION:
APPLICANT: Bertino, Joseph R.
APPLICANT: Banerjee, Debabrata
APPLICANT: Tong, Youzhi
APPLICANT: Liu-Chen, Xinyue
TITLE OF INVENTION: Mutants of Thymidylate Synthase and Uses Thereof
FILE REFERENCE: D5978
CURRENT APPLICATION NUMBER: US/09/367,007C
CURRENT FILING DATE: 1999-10-15
PRIORITY FILING DATE: PCT/US98/02145
PRIORITY FILING DATE: 1998-01-03
NUMBER OF SEQ ID NOS: 39
SEQ ID NO 6
LENGTH: 24
TYPE: DNA
ORGANISM: artificial sequence
FEATURE:
NAME/KEY: primer bind
LOCATION: 645..672
OTHER INFORMATION: Antisense primer hts-2b for PCR amplification of part
OTHER INFORMATION: of 890 base pair fragment of human TS gene
OTHER INFORMATION: from nucleotide 50 to the C-terminus
US-09-367-007C-6

Query Match 1.6%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 727 CTGTCTGCCAGCTGTACACAGAGA 750
Db 24 CTGTCTGCCAGCTGTACACAGAGA 1

RESULT 12

US-09-367-007C-8/C
; Sequence 8, Application US/09367007C
; Patent No. 6416987
; GENERAL INFORMATION:
; APPLICANT: Bertino, Joseph R.
; APPLICANT: Banerjee, Debabrata
; APPLICANT: Tong, Youzhi
; APPLICANT: Liu-Chen, Xinyue
; TITLE OF INVENTION: Mutants of Thymidylate Synthase and Uses Thereof
; FILE REFERENCE: D5978
; CURRENT APPLICATION NUMBER: US/09/367,007C
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/US98/02145
; PRIOR FILING DATE: 1998-01-03
; NUMBER OF SEQ ID NOS: 39
; SEQ ID NO 8
; LENGTH: 24
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: primer bind
; LOCATION: 939..915
; OTHER INFORMATION: Antisense primer hTS-3B for PCR amplification of part
; OTHER INFORMATION: of 890 base pair fragment of human TS gene
; OTHER INFORMATION: from nucleotide 50 to the C-terminus
US-09-367-007C-8

Query Match 1.6%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1021 ACTATTAATGGAATGCGCTGTT 1044
DB 24 ACTTAAATGGAATGCGCTGT 1

RESULT 13
US-09-367-007C-13
; Sequence 13, Application US/09367007C
; Patent No. 6416987
; GENERAL INFORMATION:
; APPLICANT: Bertino, Joseph R.
; APPLICANT: Banerjee, Debabrata
; APPLICANT: Tong, Youzhi
; APPLICANT: Liu-Chen, Xinyue
; TITLE OF INVENTION: Mutants of Thymidylate Synthase and Uses Thereof
; FILE REFERENCE: D5978
; CURRENT APPLICATION NUMBER: US/09/367,007C
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/US98/02145
; PRIOR FILING DATE: 1998-01-03
; NUMBER OF SEQ ID NOS: 39
; SEQ ID NO 13
; LENGTH: 24
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: primer bind
; LOCATION: 756..779
; OTHER INFORMATION: Sense primer hTS-6A for PCR amplification of part
; OTHER INFORMATION: of 890 base pair fragment of human TS gene
; OTHER INFORMATION: from nucleotide 50 to the C-terminus
US-09-367-007C-13

Query Match 1.6%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 861 GGGAGATGACATATTACCTGAA 884
DB 1 GGGAGATGACATATTACTGAA 24

RESULT 14
US-09-367-007C-10/C
; Sequence 10, Application US/09367007C
; Patent No. 6416987
; GENERAL INFORMATION:
; APPLICANT: Bertino, Joseph R.
; APPLICANT: Banerjee, Debabrata
; APPLICANT: Tong, Youzhi
; APPLICANT: Liu-Chen, Xinyue
; TITLE OF INVENTION: Mutants of Thymidylate Synthase and Uses Thereof
; FILE REFERENCE: D5978
; CURRENT APPLICATION NUMBER: US/09/367,007C
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/US98/02145
; PRIOR FILING DATE: 1998-01-03
; NUMBER OF SEQ ID NOS: 39
; SEQ ID NO 10
; LENGTH: 23
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: primer bind
; LOCATION: 210..188
; OTHER INFORMATION: Antisense primer hTS-4B for PCR amplification of part
; OTHER INFORMATION: of 890 base pair fragment of human TS gene
; OTHER INFORMATION: from nucleotide 50 to the C-terminus
US-09-367-007C-10

Query Match 1.5%; Score 23; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 CCGGCTACAGCCTGAGAGATGAA 315
DB 23 CCGGCTACAGCCTGAGAGATGAA 1

RESULT 15
US-09-367-007C-12/C
; Sequence 12, Application US/09367007C
; Patent No. 6416987
; GENERAL INFORMATION:
; APPLICANT: Bertino, Joseph R.
; APPLICANT: Banerjee, Debabrata
; APPLICANT: Tong, Youzhi
; APPLICANT: Liu-Chen, Xinyue
; TITLE OF INVENTION: Mutants of Thymidylate Synthase and Uses Thereof
; FILE REFERENCE: D5978
; CURRENT APPLICATION NUMBER: US/09/367,007C
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/US98/02145
; PRIOR FILING DATE: 1998-01-03
; NUMBER OF SEQ ID NOS: 39
; SEQ ID NO 12
; LENGTH: 23
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: primer bind
; LOCATION: 503..481
; OTHER INFORMATION: Antisense primer hTS-5B for PCR amplification of part
; OTHER INFORMATION: of 890 base pair fragment of human TS gene
; OTHER INFORMATION: from nucleotide 50 to the C-terminus
US-09-367-007C-12

Query Match 1.5%; Score 23; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 586 CTGCAAGATGATGACACCAT 608
DB 23 CTGCAAGATGATGACACCAT 1

Sun Dec 1 18:15:47 2002

us-09-963-333-6.oligo.rni

Page 8

Search completed: November 29, 2002, 03:30:35
Job time : 100.613 secs

SUMMARIES

SUMMARIES

1
2
3
4
5
6
7
8
9
10
11
12

AX397715 Se

X02308 Huma

BC002567 Hc

AB062290 HC

AX341901 Se

AR144965 Se

AX330682 Se

AX335755 Se

D00596 Homo

AP001178 HO

AC020697 HO

G07105 hum

AC020697 HO
Y67098 H 93

AX135281 Se
AX192876 Se

AK093873 HO

GI4563 hum

BC028301 HO

AX397648 Se

AF279905 HO

AF279907 HO

AX135264 Seq

PAT 14-

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100. 101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 840. 84

Pred. No. is the number of results predicted by chance to have a

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|---------------------|
| 1 | 1394 | 90.8 | 1536 | 6 | ARI03889 | ARI03889 Sequence |
| 2 | 1394 | 90.8 | 1536 | 6 | AX397715 | AX397715 Sequence |
| 3 | 1394 | 90.8 | 1536 | 6 | AX482020 | AX482020 Sequence |
| 4 | 1394 | 90.8 | 1536 | 6 | HSTSYN1 | X02308 Human mRNA |
| 5 | 1358 | 88.4 | 1569 | 9 | BC013919 | BC013919 Homo sapi |
| 6 | 1357 | 88.3 | 1533 | 6 | BC002567 | BC002567 Homo sapi |
| 7 | 1346 | 87.6 | 3298 | 6 | AX281791 | AX281791 Sequence |
| 8 | 1489 | 77.4 | 1567 | 9 | AB062290 | AB062290 Homo sapi |
| 9 | 942 | 61.3 | 942 | 6 | AX069365 | AX069365 Sequence |
| 10 | 942 | 61.3 | 942 | 6 | AX397717 | AX397717 Sequence |
| 11 | 532 | 34.6 | 535 | 6 | AX341901 | AX341901 Sequence |
| 12 | 490 | 31.9 | 693 | 9 | AB077207 | AB077207 Homo sapi |
| 13 | 487 | 31.7 | 18596 | 6 | ARI45465 | ARI45465 Sequence |
| 14 | 487 | 31.7 | 18596 | 6 | AX050451 | AX050451 Sequence |
| 15 | 487 | 31.7 | 18596 | 6 | AX330682 | AX330682 Sequence |
| 16 | 487 | 31.7 | 18596 | 6 | AX330906 | AX330906 Sequence |
| 17 | 487 | 31.7 | 18596 | 6 | AX335755 | AX335755 Sequence |
| 18 | 487 | 31.7 | 18596 | 6 | AX397714 | AX397714 Sequence |
| 19 | 487 | 31.7 | 18596 | 6 | AX408943 | AX408943 Sequence |
| 20 | 487 | 31.7 | 18596 | 9 | HUMTSL | D00556 Homo sapien |
| 21 | 487 | 31.7 | 43507 | 9 | AFI05057 | AFI05057 Homo sapi |
| 22 | 487 | 31.7 | 152711 | 9 | AP001178 | AP001178 Homo sapi |
| 23 | 487 | 31.7 | 159974 | 2 | AC021474 | AC021474 Homo sapi |
| 24 | 487 | 31.7 | 165742 | 2 | AC020697 | AC020697 Homo sapi |
| 25 | 456 | 29.7 | 840 | 9 | AB077208 | AB077208 Homo sapi |
| 26 | 379 | 24.7 | 849 | 11 | G07105 | G07105 human SRS |
| 27 | 297 | 19.3 | 1186 | 9 | HUMTS | D00517 Homo sapien |
| 28 | 294 | 17.8 | 165742 | 2 | AC020697 | AC020697 Homo sapi |
| 29 | 168 | 10.9 | 1817 | 9 | HSTHSYN | X67098 H. sapiens r |
| 30 | 150 | 9.8 | 196 | 6 | AX135281 | AX135281 Sequence |
| 31 | 142 | 9.2 | 346 | 6 | AX192876 | AX192876 Sequence |
| 32 | 132 | 8.6 | 150700 | 2 | AP001097 | AP001097 Homo sapi |
| 33 | 117 | 7.6 | 2388 | 9 | AK093873 | AK093873 Homo sapi |
| 34 | 111 | 7.2 | 250 | 11 | G14813 | G14813 human SRS S |
| 35 | 111 | 7.2 | 360 | 11 | G14653 | G14653 human SRS S |
| 36 | 111 | 7.2 | 1133 | 6 | AX397748 | AX397748 Sequence |
| 37 | 107 | 7.0 | 860 | 9 | BC028301 | BC028301 Homo sapi |
| 38 | 107 | 7.0 | 1131 | 6 | AX397698 | AX397698 Sequence |
| 39 | 106 | 6.9 | 1161 | 6 | AX397648 | AX397648 Sequence |
| 40 | 106 | 6.9 | 168 | 9 | AFI27520 | AFI27520 Homo sapi |
| 41 | 85 | 5.5 | 249 | 9 | AF279905 | AF279905 Homo sapi |
| 42 | 82 | 5.3 | 140 | 9 | AF127519 | AF127519 Homo sapi |
| 43 | 62 | 4.0 | 165 | 9 | AP279907 | AP279907 Homo sapi |
| 44 | 62 | 4.0 | 335 | 9 | AF134215 | AF134215 Homo sapi |
| 45 | 59 | 3.8 | 196 | 6 | AX135264 | AX135264 Sequence |

ALIGNMENTS

| | | | | |
|------------|---|-------------|--------|-----------------|
| RESULT 1 | | | | |
| ARI03889 | ARI03889 | 1536 bp | DNA | PAT 14-FEB-2000 |
| LOCUS | ARI03889 | | | |
| DEFINITION | Sequence 1 from patent US 6087489. | | linear | |
| ACCESSION | ARI03889 | | | |
| VERSION | ARI03889.1 | GI:12815477 | | |
| KEYWORDS | . | | | |
| SOURCE | Unknown. | | | |
| ORGANISM | Unknown. | | | |
| REFERENCE | Unclassified. | | | |
| AUTHORS | 1 (bases 1 to 1536) | | | |
| TITLE | Dam,N.M.
Antiense oligonucleotide modulation of human thymidylate synthase
expression | | | |
| JOURNAL | Patent: US 6087489-A 1 11-JUL-2000; | | | |

QY 361 GAGAGATTGCTGTGTTTATCAAGGATCCAAATGCTAAAGCTGTCTTCCAAAGCA 420
 DB 361 GAGAGATTGCTGTGTTTATCAAGGATCCAAATGCTAAAGCTGTCTTCCAAAGCA 420
 QY 421 GTGAAATCTGGAGTGCATGATCCGAGACTTTTGGACAGCTGGGATTTCTCCAC 480
 DB 421 GTGAAATCTGGAGTGCATGATCCGAGACTTTTGGACAGCTGGGATTTCTCCAC 480
 QY 481 AGAGAAAGAGGAGCTTGGGCTGATTTATGCTTCCAGTGGAGCACTTTGGGAGAGA 540
 DB 481 AGAGAAAGAGGAGCTTGGGCTGATTTATGCTTCCAGTGGAGCACTTTGGGAGAGA 540
 QY 541 TACAGAGTATGAGATGATTTATGAGACAGGAGTGCACCACTGCAGAAAGTATG 600
 DB 541 TACAGAGTATGAGATGATTTATGAGACAGGAGTGCACCACTGCAGAAAGTATG 600
 QY 601 GACACCATCAAAACCAACCTGACGACAGAGATCATCATGTGCTGGATCCAGAA 660
 DB 601 GACACCATCAAAACCAACCTGACGACAGAGATCATCATGTGCTGGATCCAGAA 660
 QY 661 GATCTTCTGTGATGGGCTGCTCCATGCAATGCCCTTGCAGATTCTATGTGTGAAC 720
 DB 661 GATCTTCTGTGATGGGCTGCTCCATGCAATGCCCTTGCAGATTCTATGTGTGAAC 720
 QY 721 AGTGAAGCTGTGCTGACAGCTGTACAGAGATGGGAGCATGGGCTTGGCTTTTC 780
 DB 721 AGTGAAGCTGTGCTGACAGCTGTACAGAGATGGGAGCATGGGCTTGGCTTTTC 780
 QY 781 AACATGCGCAGTACGCCCTGCTCAGTACATGATTTGGGACATCAAGGAGGAGCA 840
 DB 781 AACATGCGCAGTACGCCCTGCTCAGTACATGATTTGGGACATCAAGGAGGAGCA 840
 QY 841 GGTGATTTATACACATTTGGGAGATGACATATTTACTGATTCATCGAGCCACTG 900
 DB 841 GGTGATTTATACACATTTGGGAGATGACATATTTACTGATTCATCGAGCCACTG 900
 QY 901 AAAATTCAGCTTCAAGGAGAACCAAGCTTTCCAAAGCTCAGATTTCTTCCAAAGTT 960
 DB 901 AAAATTCAGCTTCAAGGAGAACCAAGCTTTCCAAAGCTCAGATTTCTTCCAAAGTT 960
 QY 961 GAGAAATTTGATGACTTCAAAAGTGAAGCTTCAAGTGAAGGATCAATCCGATCCA 1020
 DB 961 GAGAAATTTGATGACTTCAAAAGTGAAGCTTCAAGTGAAGGATCAATCCGATCCA 1020
 QY 1021 ACTATTAATGAGAAATGCTGTTTGAAGGCTTCAAGAGGCTNCAAGATTTGTGA 1080
 DB 1021 ACTATTAATGAGAAATGCTGTTTGAAGGCTTCAAGAGGCTTCAAGATTTGTGA 1080
 QY 1081 GTCTTTAAGGGGTGGGCTGATGCCAGGTAAGTCTTTTCTCTAAAGAAAGAAAG 1140
 DB 1081 GTCTTTAAGGGGTGGGCTGATGCCAGGTAAGTCTTTTCTCTAAAGAAAGAAAG 1140
 QY 1141 AACTAGTCAAAATCTGCGTGAACCTATCATGTTATTTAATTTTAAGATGTGGCACT 1200
 DB 1141 AACTAGTCAAAATCTGCGTGAACCTATCATGTTATTTAATTTTAAGATGTGGCACT 1200
 QY 1201 GGCAAATGATCTGTGCCAGTCTTTTCCATAATTAAGCTTTGATTTAACTCACTAGAG 1260
 DB 1201 GGCAAATGATCTGTGCCAGTCTTTTCCATAATTAAGCTTTGATTTAACTCACTAGAG 1260
 QY 1261 GTATCTGACAAATGCTGAGTTATGAACAAAGTGAAGAGATGAATGATGTGCTTTAG 1320
 DB 1261 GTATCTGACAAATGCTGAGTTATGAACAAAGTGAAGAGATGAATGATGTGCTTTAG 1320
 QY 1321 CAAGAAATGATGTGCTCAATTCACATTCCTATTAAGAGAGTTGGTGAATTTTAC 1380
 DB 1321 CAAGAAATGATGTGCTCAATTCACATTCCTATTAAGAGAGTTGGTGAATTTTAC 1380
 QY 1381 AAGCTATTTTGGATATTTTGAATATTTTAAGATTTTCAAGGATTTTCTTCAAAAT 1440
 DB 1381 AAGCTATTTTGGATATTTTGAATATTTTGAATATTTTCAAGGATTTTCTTCAAAAT 1440

QY 1441 CTGAGGAGAGCTGAGTACACCATCATCATGATGATGATGTTGATTAAGCTTTA 1496
 DB 1441 CTGAGGAGAGCTGAGTACACCATCATCATGATGATGATGATGTTGATTAAGCTTTA 1496

RESULT 4

HSTS YN1

LOCUS HSTS YN1 1536 bp mRNA linear PRI 12-SEP-1993
 DEFINITION Human mRNA for thymidylate synthase (BC 2.1.1.45).
 ACCESSION X02308
 VERSION X02308.1 GI:37478
 KEYWORDS Inverted repeat; synthetase; tandem repeat.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE Takeishi, K., Kaneda, S., Ayusawa, D., Shimizu, K., Gotoh, O. and
 Seno, T.
 Nucleotide sequence of a functional cDNA for human thymidylate
 synthase
 TITLE
 JOURNAL Nucleic Acids Res. 13 (6), 2035-2043 (1985)
 MEDLINE 85215597
 PUBMED 2987839
 COMMENT Data kindly reviewed (22-OCT-1985) by Seno T.
 FEATURES
 source
 1. 1536
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 Location/Qualifiers
 misc_feature
 14. 103
 /note="triple tandemly repeated elements"
 repeat_region
 14. 16
 /note="direct repeat 1"
 misc_feature
 35. 103
 /note="pot. stem-loop structure"
 misc_feature
 35. 69
 /note="pot. stem-loop structure"
 repeat_unit
 35. 41
 /note="inverted repeat A"
 repeat_region
 42. 44
 /note="direct repeat 1"
 misc_feature
 63. 103
 /note="pot. stem-loop structure"
 repeat_unit
 63. 69
 /note="inverted repeat A"
 repeat_region
 70. 72
 /note="direct repeat 1"
 repeat_unit
 97. 103
 /note="inverted repeat A"
 repeat_region
 104. 106
 /note="direct repeat 1"
 CDS
 106. 1047
 /note="thymidylate synthase (aa 1-313)"
 /codon_start=1
 /protein_id="CA26178.1"
 /db_xref="GI:37479"
 /translation="MPVAGSELPRRLPPLPAOEDARPRPHGLQYLIOHILRCG
 VRKDRRTGTTLVSFGQKRYSLRDEPFLTLTKRVVFPKGVLELLFPIKSTNAKELS
 SKGVYWDANGSRDFLDSIGFSTREGLGVYGFQMRHFCAETRMEDSDYSGQVYD
 LQRIVDITKTPDPRRIIMCAMPDPRDPLMALPCHALCQFYVNSLSCLQYQNSGD
 MGLGVPRIVASVALLTYMIAHTLGLKPGDFHTLGDALIVLNHIEPLKIQLOREPRPF
 PKRLIKRVEKIDDFKAEDFQLEGNPHPTIKEMAY"

misc_feature
 1231. 1236
 /note="pot. polyadenylation signal"
 misc_feature
 1519. 1524
 /note="pot. polyadenylation signal"
 polyA_site
 1536
 /note="polyadenylation site"
 BASE COUNT 390 a 369 c 399 g 378 t
 ORIGIN

Query Match

90.8%; Score 1394; DB 9; Length 1536;

GENERAL ATOMICS (US)

FEATURES Location/Qualifiers
source 1..942
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Human thymidylate synthase gene: exons 1-8"

BASE COUNT 236 a 247 c 257 g 202 t
ORIGIN

Query Match 61.3%; Score 942; DB 6; Length 942;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 942; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

106 ATGCGTGTGGCCGGCTCGAGCTGCGCGCGGCGCTTGGCCCGCCCGCGCAAGAGCGG 165
1 ATGCTGTGTGGCGGCTCGAGCTGCGCGCGGCGCTTGGCCCGCCCGCGCAAGAGCGG 60
166 GACGCCGAGCCGCTGCGCGCGGAGCTGCACTGACCTGCGGCGAGATCAACACATC 225
61 GACGCCGAGCCGCTGCGCGCGGAGCTGCACTGACCTGCGGCGAGATCAACACATC 120
226 CTCCTGCTGCGCGCTGCAAGAGAGACCGGACCGGCACTCTGTGCTATTCGCG 285
121 CTCCTGCTGCGCGCTGCAAGAGAGACCGGACCGGCACTCTGTGCTATTCGCG 180
286 ATGCAAGCGCGCTGCAAGAGAGATTCCTCTGCTGCAACCAACGCTGCTC 345
181 ATGCAAGCGCGCTGCAAGAGATTCCTCTGCTGCAACCAACGCTGCTC 240
346 TGAAGGGGTGTTTGAAGAGATTCCTGCTGCAACCAACGCTGCTC 405
241 TGAAGGGGTGTTTGAAGAGATTCCTGCTGCAACCAACGCTGCTC 300
406 CTGCTCTTCCAGAGAGATTCCTGCTGCAACCAACGCTGCTC 465
301 CTGCTCTTCCAGAGAGATTCCTGCTGCAACCAACGCTGCTC 360
466 CTGCGATTTCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 525
361 CTGCGATTTCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
526 CATTTTGGGGCAGATTCAGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 585
421 CATTTTGGGGCAGATTCAGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 480
586 CTGCAAGAGATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTCAG 645
481 CTGCAAGAGATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTCAG 540
646 GCTTGGATTCAGAGATTCCTCTGATGAGCGCTGCTGCAATGCGCTGCTGCGCAG 705
541 GCTTGGATTCAGAGATTCCTCTGATGAGCGCTGCTGCAATGCGCTGCTGCGCAG 600
706 TTCTATGTGTGAACAGTGAAGTGTCTGCTGCACTGCTGCAAGATTCAGAGATTCAG 765
601 TTCTATGTGTGAACAGTGAAGTGTCTGCTGCACTGCTGCAAGATTCAGAGATTCAG 660
766 CTGCGTGTGCTTCAACATGCGAGCTGCGCTGCTGCACTGCTGCAAGATTCAGATTC 825
661 CTGCGTGTGCTTCAACATGCGAGCTGCGCTGCTGCACTGCTGCAAGATTCAGATTC 720
826 ACGGGCTTGAAGCAGGTGATTCATCACTTTGGAGATTCAGATTCAGATTCAGATTC 885
721 ACGGGCTTGAAGCAGGTGATTCATCACTTTGGAGATTCAGATTCAGATTCAGATTC 780
886 CATATGAGGCTGAAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 945
781 CATATGAGGCTGAAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 840
946 ATTCTTGAAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 1005
841 ATTCTTGAAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 900

QY 1006 TACATCCCATCATCACTATTAAATGAAGATGCTGTTAG 1047
DB 901 TACATCCCATCATCACTATTAAATGAAGATGCTGTTAG 942

RESULT 10

AX397717 942 bp DNA linear PAT 18-MAY-2002
LOCUS AX397717
DEFINITION Sequence 143 from Patent WO0210366.
ACCESSION AX397717
VERSION AX397717.1 GI:21068438
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 Chen, H., Freimer, N.B. and Novak, T.
AUTHORS Methods and compositions for diagnosing and treating chromosome-18p
related disorders
TITLE Patent: WO 0210366-A 143 07-FEB-2002;
JOURNAL Millennium Pharmaceuticals, Inc. (US); The Regents of The
University of California (US)
LOCATION/Qualifiers

FEATURES

source 1..942
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 236 a 247 c 257 g 202 t
ORIGIN

Query Match 61.3%; Score 942; DB 6; Length 942;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 942; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

106 ATGCGTGTGGCCGGCTCGAGCTGCGCGCGGCGCTTGGCCCGCCCGCGCAAGAGCGG 165
1 ATGCTGTGTGGCGGCTCGAGCTGCGCGCGGCGCTTGGCCCGCCCGCGCAAGAGCGG 60
166 GACGCCGAGCCGCTGCGCGCGGAGCTGCACTGACCTGCGGCGAGATCAACACATC 225
61 GACGCCGAGCCGCTGCGCGCGGAGCTGCACTGACCTGCGGCGAGATCAACACATC 120
226 CTCCTGCTGCGCGCTGCAAGAGAGACCGGACCGGCACTCTGTGCTATTCGCG 285
121 CTCCTGCTGCGCGCTGCAAGAGAGACCGGACCGGCACTCTGTGCTATTCGCG 180
286 ATGCAAGCGCGCTGCAAGAGAGATTCCTCTGCTGCAACCAACGCTGCTC 345
181 ATGCAAGCGCGCTGCAAGAGATTCCTCTGCTGCAACCAACGCTGCTC 240
346 TGAAGGGGTGTTTGAAGAGATTCCTGCTGCAACCAACGCTGCTC 405
241 TGAAGGGGTGTTTGAAGAGATTCCTGCTGCAACCAACGCTGCTC 300
406 CTGCTCTTCCAGAGAGATTCCTGCTGCAACCAACGCTGCTC 465
301 CTGCTCTTCCAGAGAGATTCCTGCTGCAACCAACGCTGCTC 360
466 CTGCGATTTCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 525
361 CTGCGATTTCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
526 CATTTTGGGGCAGATTCAGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 585
421 CATTTTGGGGCAGATTCAGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 480
586 CTGCAAGAGATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTCAG 945
481 CTGCAAGAGATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTCAG 840
946 GCTTGGATTCAGAGATTCCTCTGATGAGCGCTGCTGCAATGCGCTGCTGCGCAG 705
541 GCTTGGATTCAGAGATTCCTCTGATGAGCGCTGCTGCAATGCGCTGCTGCGCAG 600

QY 706 TTCTATGTGTGAACAGTGTGCTGCTCCAGCTGTATACAGAGATCGGAGACATGAGC 765
 DB 601 TTCTATGTGTGAACAGTGTGCTGCTCCAGCTGTATACAGAGATCGGAGACATGAGC 660
 QY 766 CTCGGTGTGCTTCAACATCGCAGCTACGCCCTGTCTCACTGATGATTCGCGACATC 825
 DB 661 CTCGGTGTGCTTCAACATCGCAGCTACGCCCTGTCTCACTGATGATTCGCGACATC 720
 QY 826 ACGGGCTGAGAGCCAGGTGTGCTTATACACATTCGAGATGACATATTTACTGAT 885
 DB 721 ACGGGCTGAGAGCCAGGTGTGCTTATACACATTCGAGATGACATATTTACTGAT 780
 QY 886 CACATCGAGCCAGTGAATAATTCAGTCTACGCGAGAACCCAGATCTTCCGAAAGCTGAG 945
 DB 781 CACATCGAGCCAGTGAATAATTCAGTCTACGCGAGAACCCAGATCTTCCGAAAGCTGAG 840
 QY 946 ATTCTTGGAAAAGTTGAGAAAATTTGATGACTTCAAGCTGAGACCTTTCAGATGAGAG 1005
 DB 841 ATTCTTGGAAAAGTTGAGAAAATTTGATGACTTCAAGCTGAGACCTTTCAGATGAGAG 900
 QY 1006 TACAATCGCATCCACATTTAAATGAAATGCGCTTTAG 1047
 DB 901 TACAATCGCATCCACATTTAAATGAAATGCGCTTTAG 942
 RESULT 11
 AX341901 535 bp DNA linear PAT 10-JAN-2002
 LOCUS Sequence 2148 from Patent WO0196388.
 DEFINITION AX341901
 ACCESSION AX341901
 VERSION AX341901.1 GI:18137883
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Jiang, Y., Harlocker, S. J., and Secrist, H.
 AUTHORS Composition and methods for the therapy and diagnosis of colon
 TITLE Cancer
 JOURNAL Patent: WO 0196388-A-2148 20-DEC-2001;
 FEATURES
 CORIXA CORPORATION (US)
 source Location/Qualifiers
 1..535
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 135 a 129 c 151 g 118 t 2 others
 ORIGIN
 Query Match 34.6%; Score 532; DB 6; Length 535;
 Best Local Similarity 100.0%; Pred. No. 1.9e-308; Indels 0; Gaps 0;
 Matches 532; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 203 ACCCTGGGGAGATCCAAACATCTCCGCTGCGGCTGAGAGAGACACGCGAGCA 262
 DB 1 ACCCTGGGGAGATCCAAACATCTCCGCTGCGGCTGAGAGAGACACGCGAGCA 60
 QY 263 CCGGCAACCTGTGTGATTTGCGATGACGGCGGCTACAGCTGAGAGATGATTCCTC 322
 DB 61 CCGGCAACCTGTGTGATTTGCGATGACGGCGGCTACAGCTGAGAGATGATTCCTC 120
 QY 323 TGTGACAAACCAAGCTGTCTTCTGAAAGGATGTTTGAAGAGATGCGTGATTTCA 382
 DB 121 TGTGACAAACCAAGCTGTCTTCTGAAAGGATGTTTGAAGAGATGCGTGATTTCA 180
 QY 383 AGGATCCAGATGCTTAAAGAGCTGTCTTCAAGGAGTGAATCTGGATGCAATG 442
 DB 181 AGGATCCAGATGCTTAAAGAGCTGTCTTCAAGGAGTGAATCTGGATGCAATG 240
 QY 443 GATCCGAGACTTTTGAACAGCTGGAGATTCACACAGAGAAAGGAGACTTGGGCC 502
 DB 241 GATCCGAGACTTTTGAACAGCTGGAGATTCACACAGAGAAAGGAGACTTGGGCC 300

QY 503 CAGTTAATGGCTTCAGTGGAGGCAATTTTGGGGCAAGATATACAGATATGAGATT 562
 DB 301 CAGTTAATGGCTTCAGTGGAGGCAATTTTGGGGCAAGATATACAGATATGAGATT 360
 QY 563 ATTCAAGAGAGGAGTTGACCAACTGCAAGAGTATGATGACCAATCAAAACCTG 622
 DB 361 ATTCAAGAGAGGAGTTGACCAACTGCAAGAGTATGATGACCAATCAAAACCTG 420
 QY 623 ATCAGAGAGATATCATGTGCGCTTGAATTCAGAGATCTTCTGATGCGCTGC 682
 DB 421 ATCAGAGAGATATCATGTGCGCTTGAATTCAGAGATCTTCTGATGCGCTGC 480
 QY 683 CTCATGCAATCCCTCTGCGCAAGTTCTATGTGTGAGAACAGTACGCTGCTG 734
 DB 481 CTCATGCAATCCCTCTGCGCAAGTTCTATGTGTGAGAACAGTACGCTGCTG 532
 RESULT 12
 AB077207 693 bp mRNA linear PRI 09-JAN-2002
 LOCUS Homo sapiens TS mRNA for thymidylate synthase, complete cds,
 cell_line:MN45.
 DEFINITION AB077207
 ACCESSION AB077207
 VERSION AB077207.1 GI:18150848
 KEYWORDS
 SOURCE Homo sapiens cell_line:MN45 cDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Hisatomi, H. and Nagao, K.
 AUTHORS Homo sapiens TS mRNA for thymidylate synthase
 JOURNAL Published Only in Database (2002)
 REFERENCE
 2 (bases 1 to 693)
 AUTHORS Hisatomi, H. and Nagao, K.
 TITLE Direct Submission
 JOURNAL Submitted (07-JAN-2002) Hisashi Hisatomi, SRL, Inc., Center for
 Molecular Biology and CytoGenetics; 5-6-50 Shin-machi, Hino, Tokyo
 191-0002, Japan (E-mail: hisatomi@r1.srl-inc.co.jp,
 Tel:81426483793 (ex. 7926), Fax:81426484043)
 FEATURES
 source Location/Qualifiers
 1..693
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="18"
 /map="18p11.32"
 /cell_line="MN45"
 1..693
 /gene="TS"
 1..693
 /note="TS alternative splicing variant The protein encoded
 by this gene is 83aa shorter than the full-length isoform.
 This variant uses an alternative splice site resulting in
 it lacking 249nt in the coding region compared to the
 full-length isoform reported in Acc# D00596"
 /codon_start=1
 /evidence=experimental
 /product="thymidylate synthase"
 /protein_id="BAB83676.1"
 /db_xref="GI:18150849"
 /translation="MPTASGSELPKPPAPQEDAPRPPEHGLQYIOIHILRG
 VKMDRTGTGLSVFGQARYSILRYSGQVDQQRVITDTIKTPDPDRITMCMAMP
 DPLALPCHALCOFYVNSSEISCOLYRSGMDGVSPNIVASVALLTYMAHITGL
 RQDPIHTIGDAHYLNHLEPKIQIQREPRFPFKRLTKRKVIDFAEDFOIEGY
 NHPITIKENAV"
 BASE COUNT 171 a 201 c 182 g 139 t
 ORIGIN
 Query Match 31.9%; Score 490; DB 9; Length 693;
 Best Local Similarity 100.0%; Pred. No. 4.2e-283; Indels 0; Gaps 0;
 Matches 490; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

| Query Match | 31 %; | Score 487; | DB 6; | Length 18596; |
|---|-----------------|---------------------|-----------|---------------|
| Best Local Similarity | 99.7%; | Pred. No. 4.5e-281; | | |
| Matches 587; | Conservative 0; | Mismatches 2; | Indels 0; | Gaps 0; |
| QY 908 AGCTTCACGGAGAACCCAGACCTTTTCCCAACCTCAGATTCTTGAAAAGTGTGAGAAA | | | | 967 |
| Db 15611 AGCTTCACGGAGAACCCAGACCTTTTCCCAAGCTCAGATTCTTCCAAAAGTTGAGAAA | | | | 15670 |
| QY 968 TTGATGACTTAAAGCTGGAAGACTTTCCAGATTGAAGGTACATCCGCATCCAACTATTA | | | | 1027 |
| Db 15671 TTGATGACTTAAAGCTGGAAGACTTTCCAGATTGAAGGTACAATCCGCATCCAACTATTA | | | | 15730 |
| QY 1028 AAATGAAATGGCTGTTAGGGTCTTTAAAGAGCTGGAAGATATGTGACCTTTA | | | | 1087 |

| | | | | |
|------------|-----------------------------|--|-------------|--------|
| | RESULT 14 | | | |
| | AX050451 | | | |
| | LOCUS | AX050451 | 18596 bp. | DNA |
| | DEFINITION | Sequence 11 from Patent WO00711754. | | linear |
| | ACCESSION | AX050451 | | |
| | VERSION | AX050451.1 | GI:12226657 | |
| | KEYWORDS | . | | |
| SOURCE | ORGANISM | human. | | |
| | | Homo sapiens | | |
| | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| | | Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. | | |
| REFERENCE | TITLE | 1 (bases 1 to 18596) | | |
| AUTHORS | | Johnson,W.G. and Stemroos,R.S. | | |
| | | Methods for diagnosing, preventing, and treating developmental disorders due to a combination of genetic and environmental factors | | |
| JOURNAL | | Patent: WO 00711754-A 11 30-NOV-2000; | | |
| | | University of Medicine and Dentistry of New Jersey (US) | | |
| FEATURES | Location/Qualifiers | | | |
| source | 1..18596 | | | |
| | /organism="Homo sapiens" | | | |
| | /db_xref="taxon:9606" | | | |
| BASE COUNT | 4521 a 3991 c 4479 g 5605 t | | | |
| ORIGIN | | | | |

| | Query Match | Best Local Match | Similarity | Score | DB 6 | Length | 18596 |
|----|-------------|--|--------------|-------|------------|--------|--------|
| | Matches | 567 | Conservative | 0 | Mismatches | 2 | Indels |
| | | | | | | | Gaps |
| QY | 908 | AGCTTACGCGAAGAACCCAGACCTTTCCCAAGCTCAGATTCTTCGAAAAAGTTGAGAAAA | 91.7% | 487 | 18596 | 18596 | 0 |
| Db | 15611 | AGCTTACGCGAAGAACCCAGACCTTTCCCAAGCTCAGATTCTTCGAAAAAGTTGAGAAAA | 99.7% | 487 | 18596 | 18596 | 0 |
| QY | 968 | TTGATGACTTCAAACTGGAAGCTTCAGATTGAAGGCTACAAATCGGCATCCACTATT | 99.7% | 487 | 18596 | 18596 | 0 |
| Db | 15671 | TTGATGACTTCAAACTGGAAGCTTCAGATTGAAGGCTACAAATCGGCATCCACTATT | 99.7% | 487 | 18596 | 18596 | 0 |
| QY | 1028 | AAATGCAATGCGCTGTTTAAAGGCTCTTCAAAAGAGCTGAAAGATATTTGAGCTTTA | 99.7% | 487 | 18596 | 18596 | 0 |
| Db | 15731 | AAATGCAATGCGCTGTTTAAAGGCTCTTCAAAAGAGCTGAAAGATATTTGAGCTTTA | 99.7% | 487 | 18596 | 18596 | 0 |
| QY | 1088 | GGGCTTGGCTGGATGCGAAGTAAAGTCTTTTGTCTAAAGAAAGAACTAAG | 99.7% | 487 | 18596 | 18596 | 0 |

Db 15851 TCAAAATCTGTCCGAGACCTATCAGTTATTAATTTTAAAGATGTGCGACTGGCAAT 15910
Qy 1208 GTAACGTGCCAGTCTCTTTCATTAATTAAGGCTTTGAGTTAACTCACTGAGGGTATCTG 1267
Db 15911 GTACAGTGCAGTCTCTTTCATTAATTAAGGCTTTGAGTTAACTCACTGAGGGTATCTG 15970
Qy 1268 ACAATGCTAGGTTATGACAAAGTGAGGAGATGAATGTATGTGCTCTTAGCAAAATC 1327
Db 15971 ACAATGCTAGGTTATGACAAAGTGAGGAGATGAATGTATGTGCTCTTAGCAAAATC 16030
Qy 1328 ATGTATGTGCATTTCAATCCCAAGTACTTAATAAGAAAGTGTGATATTTCAACAAGTAT 1387
Db 16031 ATGTATGTGCATTTCAATCCCAAGTACTTAATAAGAAAGTGTGATATTTCAACAAGTAT 16090
Qy 1388 TTTTGGATATTTTGAATATTTTAAGAATTTTCAACAAGCTATCCCTCAATCTGAGGG 1447
Db 16091 TTTTGGATATTTTGAATATTTTAAGAATTTTCAACAAGCTATCCCTCAATCTGAGGG 16150
Qy 1448 AGCTGAGTAAACCAATGCATGCATGTAGTAGTGTGGTTATGAACTTTA 1496
Db 16151 AGCTGAGTAAACCAATGCATGCATGTAGTAGTGTGGTTATGAACTTTA 16199

Search completed: November 29, 2002, 02:29:06
Job time : 3980.66 secs

Search completed: November 29, 2002, 02:29:06
Job time : 3980.66 secs

| | |
|------------|--|
| RESULT | 15 |
| LOCUS | AXJ30682 |
| DEFINITION | AXJ30682 18596 bp DNA linear PAT 09-JAN-2002 |
| ACCESSION | Sequence 1191 from Patent WO0194629. |
| VERSION | AXJ30682 |
| KEYWORDS | AXJ30682.1 GI:18103660 |
| SOURCE | . |
| ORGANISM | human. |
| REFERENCE | Homo sapiens |
| AUTHORS | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 |
| TITLE | Young,P.F., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppe,D.R. and Weaver,Z.
Cancer gene determination and therapeutic screening using signature
gene sets |
| JOURNAL | Patent: WO 0194629-A 1191 13-DEC-2001; |
| FEATURES | Avalon Pharmaceuticals (US) |
| source | location/Qualifiers |
| | 1..18596 |
| | /organism="Homo sapiens" |
| | /db xref="taxon:9606" |
| BASE COUNT | 4521 a 3991 c 4479 g 5605 t |
| ORIGIN | |

| | Query Match | Best Local Match | Similarity | Score | Length |
|---------|-------------|------------------|------------|--------------------|----------|
| Matches | 587 | Conservative | 99.7% | 487 | 18596 |
| | | | | Pred. No. 4,56-281 | |
| | | | | Mismatches 2 | Indels 0 |
| | | | | Gaps 0 | |

RESULT 3
 AAS84960/c
 ID AAS84960 standard; cDNA; 1539 BP.
 XX
 AC AAS84960;
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #20764.
 XX
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 PN WO200175067-A2.
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 PI WPI: 2001-639362/73.
 DR P-PSDB; ABG20773.
 DR
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 1; SEQ ID No 20764; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1539 BP; 396 A; 379 C; 347 G; 417 T; 0 other;
 XX
 Query Match 34.9%; Score 536; DB 23; Length 1539;
 Best Local Similarity 100.0%; Pred. No. 2e-266;
 Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 106 ATGCTTGAGCGGCTGAGAGCTGCGCGCGCCCTTGCCCGCCGACAGAGAGCGG 165
 DB 1539 ATGCTTGAGCGGCTGAGAGCTGCGCGCGCCCTTGCCCGCCGACAGAGAGCGG 1480
 QY 166 GACGCGGAGCGGCTGCGCGGACGAGAGCTGAGTACCTTGAGGAGAGATCCAAACATC 225
 DB 1479 GACGCGGAGCGGCTGCGCGGACGAGAGCTGAGTACCTTGAGGAGAGATCCAAACATC 1420

QY 226 CTCGCTGCGGCGCTTCAGAGAGACGACCGGACCGGACCGGCGCCCTGCTATTGCGC 285
 DB 1419 CTCGCTGCGGCGCTTCAGAGAGACGACCGGACCGGACCGGCGCCCTGCTATTGCGC 1360
 QY 286 ATGCAGAGCGGCTTCAGAGAGATGAAATTCCTGCTGACCAACCAACGCTGCTTC 345
 DB 1359 ATGCAGAGCGGCTTCAGAGAGATGAAATTCCTGCTGACCAACCAACGCTGCTTC 1300
 QY 346 TGAAGAGGTGTTTGGAGAGAGTGTGTGTGTTATCAAGGATCCAAATGCTAAAG 405
 DB 1299 TGAAGAGGTGTTTGGAGAGAGTGTGTGTGTTATCAAGGATCCAAATGCTAAAG 1240
 QY 406 CTGCTTCGAGAGAGATGAAATCTGGAGTGCATGATCCGAGACTTTTGGACAGC 465
 DB 1239 CTGCTTCGAGAGAGATGAAATCTGGAGTGCATGATCCGAGACTTTTGGACAGC 1180
 QY 466 CTGGGATTCCTCCAGAGAGAGAGGACTTGGGCCCAATTATGCTTCCAGTGGAGG 525
 DB 1179 CTGGGATTCCTCCAGAGAGAGAGGACTTGGGCCCAATTATGCTTCCAGTGGAGG 1120
 QY 526 CATTTTGGGCGAGAAATCAAGATATGAAATGATATTCAGGACAGGAGTTGACCA 585
 DB 1119 CATTTTGGGCGAGAAATCAAGATATGAAATGATATTCAGGACAGGAGTTGACCA 1060
 QY 586 CTGCAAGAGTGAATTGACACCATCAAAACCAACCTTGACGACAGAGAAATCATCAT 641
 DB 1059 CTGCAAGAGTGAATTGACACCATCAAAACCAACCTTGACGACAGAGAAATCATCAT 1004
 RESULT 4
 ABL38559
 ID ABL38559 standard; cDNA; 535 BP.
 XX
 AC ABL38559;
 XX
 DT 08-APR-2002 (first entry)
 XX
 DE Human colon tumour antigen polynucleotide SEQ ID NO:2148.
 XX
 KM Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
 KW colon tumour metastatic antigen; diagnosis; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200196388-A2.
 XX
 PD 20-DEC-2001.
 XX
 PF 08-JUN-2001; 2001WO-US18557.
 XX
 PR 09-JUN-2000; 2000US-210899P.
 PR 20-FEB-2001; 2001US-270216P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Jiang Y, Harlocker SL, Secrist H;
 PI WPI: 2002-114514/15.
 DR
 PT Novel isolated colon tumor polynucleotide differentially expressed in
 PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
 PT useful for inhibiting development of cancer in patient -
 XX
 PS Claim 1; SEQ ID 2148; 105bp; English.
 XX
 CC ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
 CC which were isolated from human colon tumour and colon metastatic tumour
 CC cDNA libraries. (I) have cytostatic activity and can be used in vaccine
 CC production. (I) can be used for stimulating and/or expanding T cells
 CC specific for a tumour protein on contact with the T cells. They are also
 CC useful for inhibiting the development of cancer in a patient. (I) can be
 CC used as probes or primers for nucleic acid hybridisation, for preparing

CC mutant species primers, or primers for use in genetic constructions. (I)
CC can be used in the diagnosis of a colon tumour.

XX Sequence 535 BP; 135 A; 129 C; 151 G; 118 T; 2 other;

Query Match 34.6%; Score 532; DB 24; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.3e-264;
Matches 532; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 203 ACCCTGAGGAGATCCAAACATCTCTCCGCTGCGGGCTCAAGAAAGACACCGCACGGCA 262
DB 1 ACCTGGGGGACATCCAAACATCTCTCCGCTGCGGGCTCAAGAAAGACACCGCACGGCA 60
QY 263 CCGGACCTCTGCGTATTCGGCATGACGGCGGCTACACGCTTGAGAGATTAATTCCTC 322
DB 61 CCGGACCTCTGCGTATTCGGCATGACGGCGGCTACACGCTTGAGAGATTAATTCCTC 120
QY 323 TGCCTGACACCAACGCTGCTGCGAAGAGGCTTTTGAAGAGATGCTGCTGCTTATCA 382
DB 121 TGCCTGACACCAACGCTGCTGCGAAGAGGCTTTTGAAGAGATGCTGCTGCTTATCA 180
QY 383 AGGGATCCAAATGCTTAAAGCTGCTTCCAAAGGAGTGAATTCGGATGCCAATG 442
DB 181 AGGGATCCAAATGCTTAAAGCTGCTTCCAAAGGAGTGAATTCGGATGCCAATG 240
QY 443 GATCCGAGACTTTTGGACAGCTGGGATTCCTCAACAGAGAGAGGAGACTTGGGCC 502
DB 241 GATCCGAGACTTTTGGACAGCTGGGATTCCTCAACAGAGAGAGGAGACTTGGGCC 300
QY 503 CAGTTATGCTTCCAGTGGAGGCAATTTGGGGGAGAAATACAGAGATATGGAATCAGATT 562
DB 301 CAGTTATGCTTCCAGTGGAGGCAATTTGGGGGAGAAATACAGAGATATGGAATCAGATT 360
QY 563 ATTCAAGACAGGAGATTGACCAACTGCAAGAGATGATTGACACCATCAAAACCAACCTG 622
DB 361 ATTCAAGACAGGAGATTGACCAACTGCAAGAGATGATTGACACCATCAAAACCAACCTG 420
QY 623 ACACAGAGAGATCATCATGCTGCTTGGAAATCCAGAGATCTTCTCGATGGCGCTGC 682
DB 421 ACACAGAGAGATCATCATGCTGCTTGGAAATCCAGAGATCTTCTCGATGGCGCTGC 480
QY 683 CTCATGCCATGCGCTCTGCGCATCTATGCTGGAACAGTGAAGCTGCTCG 734
DB 481 CTCATGCCATGCGCTCTGCGCATCTATGCTGGAACAGTGAAGCTGCTCG 532

RESULT 5
ID AAF31109 standard; cDNA; 18596 BP.
XX AAF31109;
AC AAF31109;
XX 27-APR-2001 (first entry)
DT
XX Thymidylate synthase coding sequence.
DE
XX Analyte-binding enzyme; analyte analysis; ss.
XX
OS Homo sapiens.
XX
PN WO200102600-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18057.
XX
PR 06-JUN-1999; 99US-0347878.
PR 06-DEC-1999; 99US-0457205.
XX
PA (GEAT ) GEN ATOMICS.
XX
PI Yuan C;
XX
```

DR WPI: 2001-071583/08.
XX Assaying method, useful for prognosis and diagnosis of disease,
PT comprises contacting sample with a mutant analyte-binding enzyme and
PT detecting binding -

XX Disclosure; Page -: 187pp; English.

CC The present invention relates to a method for assaying an analyte in a
CC sample comprising: contacting the sample with a mutant analyte-binding
CC enzyme which has binding affinity for the analyte or an immediate
CC analyte enzymatic conversion product but has attenuated catalytic
CC activity; and detecting resulting binding. The method is useful in
CC monitoring biological systems/processes, or prognosis/diagnosis of
CC disease caused by imbalances of the analytes. The present sequence is
CC a coding sequence used in the present invention.
CC Note: the present sequence is not shown in the specification, but was
CC from Genbank, using information given in the specification.

XX Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;

Query Match 31.7%; Score 487; DB 22; Length 18596;
Best Local Similarity 99.7%; Pred. No. 4.6e-241;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 908 AGCTTACGCCAGAACCCCAACCTTCCCAAGCTCAGAGATTTCTGAAAGTTGAGAAA 967
DB 15611 AGCTTACGCCAGAACCCCAACCTTCCCAAGCTCAGAGATTTCTGAAAGTTGAGAAA 15670
QY 968 TTGATGACTTCAAGCTGAAGCTTTCCAGTTGAAGGATCAATCCGATCCACTATTA 1027
DB 15671 TTGATGACTTCAAGCTGAAGCTTTCCAGTTGAAGGATCAATCCGATCCACTATTA 15730
QY 1028 AAATGSAATAGGCTGTTTGGGCTTCAAGGAGCTGTAAGAGATATGTCAGCTTTA 1087
DB 15731 AAATGSAATAGGCTGTTTGGGCTTCAAGGAGCTGTAAGAGATATGTCAGCTTTA 15790
QY 1088 GGGGTGGGCTGAGTCCAGAGGTAAGATCTTTTCTCTAAGAGAAAGAACTAGG 1147
DB 15791 GGGGTGGGCTGAGTCCAGAGGTAAGATCTTTTCTCTAAGAGAAAGAAAGCTAGG 15850
QY 1148 TCAAAAATCTGTCGGTACCTATCAGTTATTTTAAAGATGTTGCGATGSCAAAT 1207
DB 15851 TCAAAAATCTGTCGGTACCTATCAGTTATTTTAAAGATGTTGCGATGSCAAAT 15910
QY 1208 GTAATGCGCAAGTCTTTCCTCAATTAATAAGCTTTGATTACTCATGAGGATCTG 1267
DB 15911 GTAATGCGCAAGTCTTTCCTCAATTAATAAGCTTTGATTACTCATGAGGATCTG 15970
QY 1268 ACAATGCTGAGGTTATGACAAAGAGAGAGAGATGAATGATGCTTGAAGAAAC 1327
DB 15971 ACAATGCTGAGGTTATGACAAAGAGAGAGAGATGAATGATGCTTGAAGAAAC 16030
QY 1328 ATGTATGCAATTCATCCACGACTTATTAAGAGAGTGTGATTTCAAGACTAT 1387
DB 16031 ATGTATGCAATTCATCCACGACTTATTAAGAGAGTGTGATTTCAAGACTAT 16090
QY 1388 TTTTGAATATTTTGAATATTTTGAATTTTGAATTTTCAAGCTATTCCTCAATCTGAGG 1447
DB 16091 TTTTGAATATTTTGAATATTTTGAATTTTGAATTTTCAAGCTATTCCTCAATCTGAGG 16150
QY 1448 AGCTGAGTAACACCATCATCATGATGATGATGATGATGATGATGATGATGATGAT 1496
DB 16151 AGCTGAGTAACACCATCATCATGATGATGATGATGATGATGATGATGATGATGAT 16199

RESULT 6
AAC91215
ID AAC91215 standard; DNA; 18596 BP.
XX AAC91215;
AC AAC91215;
XX
XX 20-MAR-2001 (first entry)
DT
```

```

XX DE Human thymidylate synthase gene SEQ ID NO: 11.
XX KW Human; schizophrenia; developmental disorder; spina bifida cystica;
XX KW Tourette's syndrome; bipolar illness; autism; conduct disorder;
XX KW attention deficit disorder; obsessive compulsive disorder;
XX KW chronic multiple tic syndrome; learning disorder; polymorphism; ds.
XX OS Homo sapiens.
XX PN WO200071754-A1.
XX PD 30-NOV-2000.
XX PF 24-MAY-2000; 2000NO-US14354.
XX PR 25-MAY-1999; 99US-0318448.
XX PA (UYME-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
XX PI Johnson WG, Steenroos ES;
XX DR WPI; 2001-025174/03.
XX PT Diagnosing a developmental disorder, e.g. schizophrenia, by forming
XX PT datasets (DS) of genetic (e.g. genotypes of folate metabolism alleles)
XX PT and environmental variables affecting an individual and then comparing
XX PT these DS with reference DS -
XX PS Disclosure; Page 125-131, 156pp; English.
XX CC The present invention provides a novel method of estimating the
XX CC susceptibility of an individual to a developmental disorder using genetic
XX CC and environmental variables. The method can be used in the diagnosis,
XX CC prevention and treatment of disorders such as schizophrenia, spina bifida
XX CC cystica, Tourette's syndrome, bipolar illness, autism, conduct disorders,
XX CC attention deficit disorder, obsessive compulsive disorder, chronic
XX CC multiple tic syndrome and learning disorders such as dyslexia.
XX SQ Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;

Query Match 31.7%; Score 487; DB 22; Length 18596;
Best Local Similarity 99.7%; Pred. No. 4.6e-241;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 908 AGCTTCACCGAAGCCGAGACCTTTCCAAAGCTCAGGATTTCTGAAAAGTTGAGAAA 967
DB 15611 AGCTTCACCGAAGCCGAGACCTTTCCAAAGCTCAGGATTTCTGAAAAGTTGAGAAA 15670
QY 968 TTGATGATTCGAAGCTGGAAGCTTCAGATTCGAGGAGTACATCCGATCCACTATTA 1027
DB 15671 TTGATGATTCGAAGCTGGAAGCTTCAGATTCGAGGAGTACATCCGATCCACTATTA 15730
QY 1028 AATGGAATGCGCTGTTAGGCTTTTAAAGAGCTGAGGATATTCAGTCTTTA 1087
DB 15731 AATGGAATGCGCTGTTAGGCTTTTAAAGAGCTGAGGATATTCAGTCTTTA 15790
QY 1088 GGGGTTGGGCTGATGCGGAGTAAAGTTCTTTTGTCTTAAAGAAAGAGACTAGG 1147
DB 15791 GGGGTTGGGCTGATGCGGAGTAAAGTTCTTTTGTCTTAAAGAAAGAGACTAGG 15850
QY 1148 TCAAAATCTGTCCGTGACTATCAGTATTAATTTTAAAGATGTTGCCACTGCGCAAT 1207
DB 15851 TCAAAATCTGTCCGTGACTATCAGTATTAATTTTAAAGATGTTGCCACTGCGCAAT 15910
QY 1208 GTAATCTGCGAGTTCTTTCCATTAATAAAGCTTTGAGTAACTACCTGAGGATCTTG 1267
DB 15911 GTAATCTGCGAGTTCTTTCCATTAATAAAGCTTTGAGTAACTACCTGAGGATCTTG 15970
QY 1268 ACAAAGCTGAGGTTATGACAAAGTGAAGAAATGAATGATGCTCTTACCAAAAAC 1327
DB 15971 ACAAAGCTGAGGTTATGACAAAGTGAAGAAATGAATGATGCTCTTACCAAAAAC 16030

QY 1328 ATGTATGTCATTTTCATCCCAAGTACTTAAAGAGTTGGTGAATTTCAACAGTAT 1387
DB 16031 ATGTATGTCATTTTCATCCCAAGTACTTAAAGAGTTGGTGAATTTCAACAGTAT 16090
QY 1388 TTTTGAATATTTTGAATATTTTAAAGATTTTCAAGGTTATCCCTCAATCTGAGG 1447
DB 16091 TTTTGAATATTTTGAATATTTTAAAGATTTTCAAGGTTATCCCTCAATCTGAGG 16150
QY 1448 AGCTGAGTAAACACATGATCATGATGATGATGATGATGATGATGATGATGAT 1496
DB 16151 AGCTGAGTAAACACATGATCATGATGATGATGATGATGATGATGATGATGAT 16199

RESULT 7
ABN95092
ID ABN95092 standard; DNA; 18596 BP.
AC ABN95092;
XX 13-AUG-2002 (first entry)
DT 13-AUG-2002 (first entry)
XX DE Gene #1590 used to diagnose liver cancer.
XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX KW metastatic liver tumour; cytostatic; expression profile; disease state;
XX KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX OS Homo sapiens.
XX PN WO200229103-A2.
XX PD 11-APR-2002.
XX PF 02-OCT-2001; 2001WO-US30589.
XX PR 02-OCT-2000; 2000US-237054P.
XX PA (GENE-) GENE LOGIC INC.
XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX DR WPI; 2002-426119/45.
XX PT Diagnosing and detecting the progression of liver cancer,
XX PT hepatocellular carcinoma or metastatic liver tumor in a patient,
XX PT involves detecting the level of expression of two or more genes in a
XX PT liver tissue sample -
XX PS Claim 1; SEQ ID NO 1590; 298bp; English.
XX CC The invention relates to a novel method for diagnosing and detecting the
XX CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX CC tumour in a patient, and differentiating metastatic liver cancer from
XX CC hepatocellular carcinoma in a patient, involving detecting the level of
XX CC expression of two or more genes represented in ABN9503-ABN97455 in a
XX CC tissue sample. The method of the invention has hepatotropic, and
XX CC cytotropic activity. The method is useful for diagnosing and detecting
XX CC the progression of liver cancer, hepatocellular carcinoma and metastatic
XX CC liver carcinoma in a patient. The method is useful for identifying
XX CC expression profiles which serve as useful diagnostic markers as well as
XX CC markers that can be used to monitor disease states, disease progression,
XX CC drug toxicity, drug efficacy and drug metabolism.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pat_sequences.

Query Match 31.7%; Score 487; DB 24; Length 18596;
Best Local Similarity 99.7%; Pred. No. 4.6e-241;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 908 AGCTTCACCGAAGCCGAGACCTTTCCAAAGCTCAGGATTTCTGAAAAGTTGAGAAA 967

```


| | | |
|-----------|---|--------------------------------------|
| | XX | Claim 1; SEQ ID 1191; 44pp; English. |
| PT | Screening for anti-neoplastic agent involves exposing cells to a | - |
| PT | chemical agent to be tested for anti-neoplastic activity, and | |
| PT | determining a change in expression of a gene of a signature gene set | |
| PS | | |
| XX | | |
| CC | The present invention describes a method (M1) for screening for an | |
| CC | anti-neoplastic agent. The method involves exposing cells to a chemical | |
| CC | agent to be tested for anti-neoplastic activity, determining a change in | |
| CC | expression of at least one gene (I) of a signature gene set, where (II) | |
| CC | comprises a sequence (S) selected from 8447 sequences [given in AB61664 | |
| CC | to AB61110], or is at least 95% identical to (S), where a change in | |
| CC | expression is indicative of anti-neoplastic activity. (I) has cytostatic | |
| CC | activity and can be used in gene therapy. M1 can be used for screening | |
| CC | an anti-neoplastic agent, and can be used for producing a product which | |
| CC | is the data collected with respect to the anti-neoplastic agent as a | |
| CC | result of M1, and the data is sufficient to convey the chemical | |
| CC | structure and/or properties of the agent. M1 can be used in the | |
| CC | treatment of cancer such as colon, breast, stomach, lung, thyroid, | |
| CC | oesophageal, ovarian, kidney, prostate or pancreatic cancer, | |
| CC | adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, | |
| CC | infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine | |
| CC | carcinoma, papillary carcinoma and Wilms' tumour. | |
| SQ | Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other; | |
| | Query Match 31.7%; Score 487; DB 24; Length 18596; | |
| | Best Local Similarity 99.7%; Pred. No. 4.6e-241; | |
| | Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0; | |
| QY | 908 AGCTTGACGAGAACCCACAAGCTTTGCCAAGACTCAGATTCTTGAAAAGTGAAGAAA | 967 |
| DB | 15611 AGCTTAGGGAAACCAGACCTTTGCCAAGACTCGAGAATCTTGAAAAAGTGAAGAAA | 15670 |
| QY | 968 TTGTATCACTTCMAAGCTGAAGACTTCAGATTGAAGGGTCAATTCGCATCAAATA | 1027 |
| DB | 15671 TTGTATCACTTCMAAGCTGAAGACTTCAGATTGAAGGGTCAATTCGCATCAAATA | 15730 |
| QY | 1028 AAATGSAATGGCTGTTTTAGGGTGCCTTTCAAAAGAGCTNAGAGATTTTGCAATCTTA | 1087 |
| DB | 15731 AAATGSAATGGCTGTTTTAGGGTGCCTTTCAAAAGAGAGCTCGAAGGAAATTTGCAGCTTAA | 15790 |
| QY | 1088 GGGGTTGGGCTGAGTCCGAGGTAAGAAAGTCTTTTCTCTAAAGAAANAGAAAGAACTAG | 1147 |
| DB | 15791 GGGGTTGGGCTGAGTCCGAGGTAAGAAAGTCTTTTCTCTAAAGAAANAGAAAGAACTAG | 15850 |
| QY | 1148 TC AAAAATCTGTCGGRACCTATCATAGTTATATTTTAAGATGTCGCATGGCAAT | 1207 |
| DB | 15851 TC AAAAATCTGTCGGRACCTATCATAGTTATATTTTAAGATGTCGCATGGCAAT | 15910 |
| QY | 1208 GTATAGTGTCCAGATTTCTTTCATATATAAAGAGCTTTAGATTAATCACTGAGGGTATCTG | 1267 |
| DB | 15911 GTATAGTGTCCAGATTTCTTTCATATATAAAGAGCTTTAGATTAATCACTGAGGGTATCTG | 15970 |
| QY | 1268 ACAATCTGAGGTATGATGAACAAAGTAGAGAGATGAATGTATGTCTCTTAGCAAAAAC | 1327 |
| DB | 15971 ACAATCTGAGGTATGATGAACAAAGTAGAGAGATGAATGTATGTCTCTTAGCAAAAAC | 16030 |
| QY | 1328 ATGTATGTCAATTCATCCOACGTAATTATAAAGAGTTGTGTAAATTCACAAGCTAT | 1387 |
| DB | 16031 ATGTATGTCAATTCATCCOACGTAATTATAAAGAGTTGTGTAAATTCACAAGCTAT | 16090 |
| QY | 1388 TTTTGGATAATTTTATGAATRTTTTAAGAAATTCACAAGCTATTCCTGTAATCTGAGGG | 1447 |
| DB | 16091 TTTTGGATAATTTTATGAATRTTTTAAATTTTCAAGAGCTATTCCTCAATCTGAGGG | 16150 |
| QY | 1448 AGCTGAGTAAACCATGATCATGATGTGAGTGGTGTGTAAGAACTTA | 1496 |
| DB | 16151 AGCTGAGTAAACCATGATCATGATGTGAGTGGTGTGTAAGAACTTA | 16199 |
| RESULT_10 | | |
| ABL65078 | | |

ID ABL63078 standard; DNA; 18596 BP.
XX
AC ABL63078;
XX
DT 15-MAY-2002 (first entry)
XX
DE Breast cancer related gene sequence SEQ ID NO:1415.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW Stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236851P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237284P.
PR 02-OCT-2000; 2000US-237285P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237455P.
PR 03-OCT-2000; 2000US-237582P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PG, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,
PI Sopot DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity; and

PT determining a change in expression of a gene of a signature gene set -
XX
XX Claim 1; SEQ ID 1415; 44bp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer.
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX
SQ Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;
Query Match 31.7%; Score 487; DB 24; Length 18596;
Best Local Similarity 99.7%; Pred. No. 4.6e-241;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 908 AGCTGAGCGAGAACCCGAGCCCTTCCGAGCTCAGAGATTCTTGAAAGTGGAGAAA 967
DB 15611 AGCTTACGCGAGAACCCGAGCCCTTCCGAGCTCAGAGATTCTTGAAAGTGGAGAAA 15670
QY 968 TTGATGACTTCAAGCTGAGAACTTTCAGATTGAGAGGTACATCCGATCCATATTA 1027
DB 15671 TTGATGACTTCAAGCTGAGAACTTTCAGATTGAGAGGTACATCCGATCCATATTA 15730
QY 1028 AATGGAATATGCTGTTTAAAGGTCTTCAAGAGCTGAGAGATATGTGACTTTTA 1087
DB 15731 AATGGAATATGCTGTTTAAAGGTCTTCAAGAGCTGAGAGATATGTGACTTTTA 15790
QY 1088 GGGGTGGGCTGAGTCCGAGGTAAGTCTTTTCTCTTAAAGAAAGAACTGAG 1147
DB 15791 GGGGTGGGCTGAGTCCGAGGTAAGTCTTTTCTCTTAAAGAAAGAACTGAG 15850
QY 1148 TCAAAATCTGCTCCGACCTATGCTATTAATTTTGAAGTGTGCGACTGGCAAT 1207
DB 15851 TCAAAATCTGCTCCGACCTATGCTATTAATTTTGAAGTGTGCGACTGGCAAT 15910
QY 1208 GTAATGCGAGTCTTCCATATAAAGGCTTGAATTACTCAGTGAAGGTAATCG 1267
DB 15911 GTAATGCGAGTCTTCCATATAAAGGCTTGAATTACTCAGTGAAGGTAATCG 15970
QY 1268 ACAATGCTGAGTTATGAACAAAGTGAAGATGATATGCTCTTGAAGAAAAC 1327
DB 15971 ACAATGCTGAGTTATGAACAAAGTGAAGATGATATGCTCTTGAAGAAAAC 16030
QY 1328 ATGATGTCATTTCAATCCACGTAATTAAGAAAGTGGTGAATTTCAAGACTAT 1387
DB 16031 ATGATGTCATTTCAATCCACGTAATTAAGAAAGTGGTGAATTTCAAGACTAT 16090
QY 1388 TTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAAT 1447
DB 16091 TTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAAT 16150
QY 1448 AGCTGAGTAACACATGATCATGATGATGATGATGATGATGATGATGATGATGAT 1496
DB 16151 AGCTGAGTAACACATGATCATGATGATGATGATGATGATGATGATGATGATGAT 16199
RESULT 11
ID ABL67927
XX ABL67927 standard; DNA; 18596 BP.
XX

AC ABL67927;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Ovary cancer related gene sequence SEQ ID NO:6264.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN M0200194623-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US10838.
 XX
 PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-231133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237299P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX

PS Claim 1, SEQ ID 6264; 44P; English.
 XX
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 847 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms tumour.
 XX
 SQ Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;
 XX
 Query Match 31.7%; Score 487; DB 24; Length 18596;
 Best Local Similarity 99.7%; Pred. No. 4,6e-241;
 Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 XX
 QY 908 AGCTTGAAGGAAACCCAGACCTTCCCAAGCTCAGGATCTTCGAAAGTGAAGAAA 967
 DB 15611 AGCTTGAAGGAAACCCAGACCTTCCCAAGCTCAGGATCTTCGAAAGTGAAGAAA 15670
 QY 968 TTGATGACTTCAAGCTGAAGACTTTCAGATTGAAGGGTACATCCGATCACTATT 1027
 DB 15671 TTGATGACTTCAAGCTGAAGACTTTCAGATTGAAGGGTACATCCGATCACTATT 15730
 QY 1028 AATGGAATGAGCTGTATGAGGCTTTCAGAGAGCTGGAAGGATTTGCTTTA 1087
 DB 15731 AATGGAATGAGCTGTATGAGGCTTTCAGAGAGCTGGAAGGATTTGCTTTA 15790
 QY 1088 GGGGTTGGGCTGGATGCCAGGTAAGAAAGTTCTTTGCTTAAAGAAAGAACTTAAG 1147
 DB 15791 GGGGTTGGGCTGGATGCCAGGTAAGAAAGTTCTTTGCTTAAAGAAAGAACTTAAG 15850
 QY 1148 TCAGAAATCTGTCGCTGACCTTACGATTTATTTTAAAGATGGTCCACGCGAAT 1207
 DB 15851 TCAGAAATCTGTCGCTGACCTTACGATTTATTTTAAAGATGGTCCACGCGAAT 15910
 QY 1208 GTACCTGTGCGAGTTCTTTCATATATTAAGGCTTGAATCTCACTGAGGGTATCTG 1267
 DB 15911 GTACCTGTGCGAGTTCTTTCATATATTAAGGCTTGAATCTCACTGAGGGTATCTG 15970
 QY 1268 ACAATGCTGAGGTATGAACAAAGTGAAGAAATGATGCTTACGAAATAAC 1327
 DB 15971 ACAATGCTGAGGTATGAACAAAGTGAAGAAATGATGCTTACGAAATAAC 16030
 QY 1328 ATGATGATGATTTCAATCCAGCTACTTAAAGAGGTTGGTGAATTTCAACACTAT 1387
 DB 16031 ATGATGATGATTTCAATCCAGCTACTTAAAGAGGTTGGTGAATTTCAACACTAT 16090
 QY 1388 TTTTGGAAATTTTAAAGATTTTAAAGAAATTTCAACAGCTTCTCCTCAATCTGAGG 1447
 DB 16091 TTTTGGAAATTTTAAAGATTTTAAAGAAATTTCAACAGCTTCTCCTCAATCTGAGG 16150
 QY 1448 AGCTGAGTAAACCAATCATATGATGATGATGATGATGATGATGATGATGATGAT 1496
 DB 16151 AGCTGAGTAAACCAATCATATGATGATGATGATGATGATGATGATGATGATGAT 16199
 RESULT 12
 ID ABA93401/c
 XX ABA93401 standard; DNA; 45716 BP.
 AC ABA93401;
 XX

DT 22-APR-2002 (first entry)

XX Human rTS-alpha gene SEQ ID NO:5.

XX Human; rTS-alpha; rTS-beta; rTS; thymidylate synthase; chromosome 18;

KW mutant thymidylate synthase; enzyme; antitumor; antidepressant;

KW neuroleptic; neurotropic; tranquilizer; gene therapy; hypomania;

KW neuropsychiatric disorder; bipolar affective disorder; schizophrenia;

KW severe bipolar affective disorder; bipolar affective disorder;

KW major depression; attention deficit disorder; schizoaffective disorder;

KW gene; ds.

XX Homo sapiens.

OS US6323244-B1.

PN 27-NOV-2001.

PD 05-NOV-1997; 97US-0965048.

PF 05-NOV-1997; 97US-0965048.

PR 05-NOV-1997; 97US-0965048.

XX (MILL-) MILLENNIUM PHARM INC.

PA (RBCG) UNIV CALIFORNIA.

XX Chen H, Freimer NB;

PI WPI; 2002-112960/15.

DR WPI; 2002-112960/15.

XX Treating an rTS (mutant form of thymidylate synthase) - mediated

PT neuropsychiatric disorder, e.g., bipolar affective disorder, comprises

PT administering a compound, identified by an rTS assay, that interferes

PT with rTS gene product -

XX Disclosure; Fig 3A-S; 110pp; English.

XX The present invention describes a method for treating an rTS (a mutant

CC form of thymidylate synthase) mediated neuropsychiatric disorder. The

CC method comprises identifying a compound (C) that can be used to treat

CC the disorder, by contacting a small molecule test compound (I) with an

CC rTS protein (II) comprising a sequence of either 361 or 416 amino acids

CC (AB05587 or AB05588) determining whether (I) binds to (II). (C) has

CC antitumor, antidepressant, neuroleptic, neurotropic and tranquilizer

CC activities. rTS polynucleotide sequences can be used in gene therapy.

CC The method is useful in identifying compounds for treating rTS-mediated

CC neuropsychiatric disorders including bipolar affective disorder,

CC e.g., severe bipolar affective (mood) disorder, bipolar affective (mood)

CC disorder with hypomania, major depression, schizophrenia, attention

CC deficit disorder and schizoaffective disorder. The present sequence

CC represents human rTS-alpha gene which is used in the exemplification of

CC the present invention. rTS has been located to chromosome 18.

XX Sequence 45716 BP; 12099 A; 10402 C; 10435 G; 12718 T; 62 other;

SQ

Query Match 31.7%; Score 487; DB 24; Length 45716;

Best Local Similarity 99.7%; Pred. No. 4.8e-241;

Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 908 AGCTTCAGCGAGAACCCAGACCTTCCCAAGCTCAGATTCCTGCAAAAGTGTGAGAAA 967

DB 44582 AGCTTCAGCGAGAACCCAGACCTTCCCAAGCTCAGATTCCTGCAAAAGTGTGAGAAA 44523

QY 968 TTGATGACTTCAAGCTGAGACTTTCAGATTGGAAGGTTCATCCGATCCCAATATTA 1027

DB 44522 TTGATGACTTCAAGCTGAGACTTTCAGATTGGAAGGTTCATCCGATCCCAATATTA 44463

QY 1028 AAATGGAATGCTGTTTAAAGGTCTTTCAAAGAGCTGAGATTTGTCACTTTTA 1087

DB 44462 AAATGGAATGCTGTTTAAAGGTCTTTCAAAGAGCTGAGATTTGTCACTTTTA 44403

QY 1088 GGGGTGGGCTGATGCCAGGTAAGTCTTTTCTTTAAAGAAAGAAAGAACTG 1147

DB 44402 GGGGTGGGCTGATGCCAGGTAAGTCTTTTCTTTAAAGAAAGAAAGAACTG 44343

QY 1148 TCAAAATCTGTCGGTGAACCTATCACTTATTAATTTTAAAGATGTGCCATGCAAT 1207

DB 44342 TCAAAATCTGTCGGTGAACCTATCACTTATTAATTTTAAAGATGTGCCATGCAAT 44283

QY 1208 GTAACCTGCGCAGTCTTTTCCAAATTAAGAGCTTTGATTAAGTCACTGAGGTAATCG 1267

DB 44282 GTAACCTGCGCAGTCTTTTCCAAATTAAGAGCTTTGATTAAGTCACTGAGGTAATCG 44223

QY 1268 ACAATGCTGAGTATTAACAAGGTGAAGAAATGAATGTATGTCTTTCAGAAAAC 1327

DB 44222 ACAATGCTGAGTATTAACAAGGTGAAGAAATGAATGTATGTCTTTCAGAAAAC 44163

QY 1328 ATGTAATGCTGATTTCAATCCAGTACTATTAATAAGAGTGGGAATTTCCAGACTAT 1387

DB 44162 ATGTAATGCTGATTTCAATCCAGTACTATTAATAAGAGTGGGAATTTCCAGACTAT 44103

QY 1388 TTTTGGAATATTTTAAATATTTTAAAGATTTTCAAGCTATCCCTCAATCTGAGG 1447

DB 44102 TTTTGGAATATTTTAAATATTTTAAAGATTTTCAAGCTATCCCTCAATCTGAGG 44043

QY 1448 AGCTGAGTACACCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1496

DB 44042 AGCTGAGTACACCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 43994

RESULT 13

ABA93402/C

ID ABA93402 standard; DNA; 45989 BP.

XX ABA93402;

AC ABA93402;

XX 22-APR-2002 (first entry)

DT Human rTS-beta gene SEQ ID NO:6.

XX Human; rTS-beta; rTS; thymidylate synthase; chromosome 18;

KW mutant thymidylate synthase; enzyme; antitumor; antidepressant;

KW neuroleptic; neurotropic; tranquilizer; gene therapy; hypomania;

KW neuropsychiatric disorder; bipolar affective disorder; schizophrenia;

KW severe bipolar affective disorder; bipolar affective disorder;

KW major depression; attention deficit disorder; schizoaffective disorder;

KW gene; ds.

XX Homo sapiens.

OS US6323244-B1.

PN 27-NOV-2001.

PD 05-NOV-1997; 97US-0965048.

PF 05-NOV-1997; 97US-0965048.

PR 05-NOV-1997; 97US-0965048.

XX (MILL-) MILLENNIUM PHARM INC.

PA (RBCG) UNIV CALIFORNIA.

XX Chen H, Freimer NB;

PI WPI; 2002-112960/15.

DR WPI; 2002-112960/15.

XX Treating an rTS (mutant form of thymidylate synthase) - mediated

PT neuropsychiatric disorder, e.g., bipolar affective disorder, comprises

PT administering a compound, identified by an rTS assay, that interferes

PT with rTS gene product -

XX Disclosure; Fig 4A-S; 110pp; English.

XX The present invention describes a method for treating an rTS (a mutant

CC form of thymidylate synthase) mediated neuropsychiatric disorder. The

CC method comprises identifying a compound (C) that can be used to treat

CC the disorder, by contacting a small molecule test compound (I) with an

CC rTS protein (II) comprising a sequence of either 361 or 416 amino acids

CC (AB05587 or AB05588) determining whether (I) binds to (II). (C) has
 CC antianic, antidepressant, neuroleptic, nootropic and tranquilizer
 CC activities. rTS polynucleotide sequences can be used in gene therapy.
 CC The method is useful in identifying compounds for treating rTS-mediated
 CC neuropsychiatric disorders including bipolar affective disorder.
 CC e.g., severe bipolar affective (mood) disorder, bipolar affective (mood)
 CC disorder with hypomania, major depression, schizophrenia, attention
 CC deficit disorder and schizoaffective disorder. The present sequence
 CC represents human rTS-beta gene which is used in the exemplification of
 CC the present invention. rTS has been located to chromosome 18.

XX Sequence 45989 BP; 12183 A; 10457 C; 10488 G; 12799 T; 62 other;

Query Match 31.7%; Score 487; DB 24; Length 45989;
 Best Local Similarity 99.7%; Pred. No. 4.8e-241;
 Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 908 AGCTTCAGCGAAGACCCGACCTTTCCAAAGCTCAGATTCTCGAAAAGTTGAGAAA 967
 DB 44855 AGCTTCAGCGAAGACCCGACCTTTCCAAAGCTCAGATTCTCGAAAAGTTGAGAAA 44796
 QY 968 TTGATGACTTCAAAGCTGAAAGACTTCAGATTGAAAGGTACATCCGATCCACTATTA 1027
 DB 44795 TTGATGACTTCAAAGCTGAAAGACTTCAGATTGAAAGGTACATCCGATCCACTATTA 44736
 QY 1028 AATGGAATGCGCTGTTTAAAGGCTTTCGAAAGGCTGAAAGGATTTGTCAGCTTTTA 1087
 DB 44735 AATGGAATGCGCTGTTTAAAGGCTTTCGAAAGGCTGAAAGGATTTGTCAGCTTTTA 44676
 QY 1088 GGGGTGGGCTGAGTGGCGAGGTAAAGTCTTTTGTCTAAAGAGAAAGAACTAGG 1147
 DB 44675 GGGGTGGGCTGAGTGGCGAGGTAAAGTCTTTTGTCTAAAGAGAAAGAACTAGG 44616
 QY 1148 TCAAAATCTGTCCGATGACCTATCAGTTATTAATTTTAAAGATGTTGCCACTGCAAT 1207
 DB 44615 TCAAAATCTGTCCGATGACCTATCAGTTATTAATTTTAAAGATGTTGCCACTGCAAT 44556
 QY 1208 GTACTGTGCGAGTCTTTTCATTAATAAAGCTTTGAGTTACTCATGAGGATTCG 1267
 DB 44555 GTACTGTGCGAGTCTTTTCATTAATAAAGCTTTGAGTTACTCATGAGGATTCG 44496
 QY 1268 ACAATGCTGAGTTATGAACAAAGTGAAGAAATGAATGATGCTCTGCAAAAAC 1327
 DB 44495 ACAATGCTGAGTTATGAACAAAGTGAAGAAATGAATGATGCTCTGCAAAAAC 44436
 QY 1328 ATGTATGTGCAATTCATCCACGATCTTAATAAGAGTGTGAAATTTCAAGACTAT 1387
 DB 44435 ATGTATGTGCAATTCATCCACGATCTTAATAAGAGTGTGAAATTTCAAGACTAT 44376
 QY 1388 TTTTGAATATTTTAAAGATTTTAAAGATTTCAAGATTTCCCTCAATCTGAGG 1447
 DB 44375 TTTTGAATATTTTAAAGATTTTAAAGATTTCAAGATTTCCCTCAATCTGAGG 44316
 QY 1448 AGCTGATGAACACCATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1496
 DB 44315 AGCTGATGAACACCATCATGATGATGATGATGATGATGATGATGATGATGATGAT 44267

RESULT 14
 AAX24270
 ID AAX24270 standard; DNA; 566 BP.
 AC AAX24270;
 XX
 DT 02-JUL-1999 (first entry)
 XX
 DE Human thymidylate synthase DNA.
 XX
 KW Thymidylate synthase; antisense oligonucleotide; cytostatic;
 KW tumor cell; anticancer drug; treatment; cancer; antiproliferative;
 KW antioestrogen; progestogen; antiandrogen; testosterone inhibitor;
 KW anti-invasion agent; growth factor inhibitor; antimetabolite;
 KW antibiotic; alkylating agent; antimetabolic agent; radiotherapy;

KW topoisomerase inhibitor; ss.
 XX Homo sapiens.
 OS
 XX MO9915648-A1.
 PN
 XX 01-APR-1999.
 PD
 XX 17-SEP-1998; 98MO-GB02820.
 PF
 XX 06-JUN-1998; 98GB-0012140.
 PR 23-SEP-1997; 97GB-0020107.
 PR 17-OCT-1997; 97GB-0022012.
 XX
 XX (ISIS-) ISIS PHARM INC.
 PA (ZENR) ZENECA LTD.
 XX
 PI Dean NM, Koropatnick DJ, Vincent MD;
 XX
 XX WPI; 1999-254708/21.
 DR
 XX
 XX
 PS
 XX
 PS Disclosure; Fig 7, 53pp; English.
 CC This invention describes novel antisense oligonucleotides targeted to
 CC sequences in the 3' end of thymidylate synthase (TS) mRNA. Such
 CC oligonucleotides are cytostatic on their own when administered to
 CC human tumor cell lines, and also enhance the toxicity of anticancer
 CC drugs such as Tomudex administered to those cells. In addition, antisense
 CC oligonucleotides targeted to 5' sequences induce TS gene transcription.
 CC The antisense oligonucleotides are used in a method for the
 CC treatment of cancer (or a method for providing antiproliferative effect)
 CC The antisense oligonucleotides are also used in the production of a
 CC medicament for the treatment of cancer, either separately or in conjunction
 CC with a therapeutic agent such as thymidylate synthase (TS) inhibitors
 CC (e.g. Tomudex, Zeneca development compound ZD9331 etc.); cytostatic
 CC agents (e.g. antioestrogens, (anti)progestogens, antiandrogens,
 CC testosterone inhibitors, anti-invasion agents, growth factor inhibitors,
 CC etc.); antiproliferative/antineoplastic agents (e.g. antimetabolites,
 CC antitumor antibiotics, alkylating agents, antimetabolic agents,
 CC topoisomerase inhibitors, etc.); or radiotherapy.
 CC
 XX
 SQ Sequence 566 BP; 156 A; 101 C; 135 G; 174 T; 0 other;

Query Match 24.7%; Score 379; DB 20; Length 566;
 Best Local Similarity 99.8%; Pred. No. 2.8e-185;
 Matches 429; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1067 GAAGATATTTGCACTTTTAAAGGCTTGGCTGATGCCGAGTAAAGTCTTTTGGT 1126
 DB 97 GAAGATATTTGCACTTTTAAAGGCTTGGCTGATGCCGAGTAAAGTCTTTTGGT 156
 QY 1127 CTAAAGAGAAAGAACTAGGTCAAAAATCTGTCCGATGACCTTACATGTTTAAATTTTA 1186
 DB 157 CTAAAGAGAAAGAACTAGGTCAAAAATCTGTCCGATGACCTTACATGTTTAAATTTTA 216
 QY 1187 AGGATGTGCGACCTGCGAAATGTAACTGTGCAAGTCTTTCATTAATAAAGCTTTGAG 1246
 DB 217 AGGATGTGCGACCTGCGAAATGTAACTGTGCAAGTCTTTCATTAATAAAGCTTTGAG 276
 QY 1247 TTAACCTAGTGAAGGTATCTGACAAATGCTGAGTTATGAACAAAGTGAAGGATGAAT 1306
 DB 277 TTAACCTAGTGAAGGTATCTGACAAATGCTGAGTTATGAACAAAGTGAAGGATGAAT 336
 QY 1307 GTATGTGCTTTAGCAAAAACATGATGTGCAATTCATCCGATGATCTTAATAAAGG 1366
 DB 337 GTATGTGCTTTAGCAAAAACATGATGTGCAATTCATCCGATGATCTTAATAAAGG 396
 QY 1367 TTGATGAATTTACAAAGTATTTTGAATATTTTAAAGATTTTAAAGATTTTCAAG 1426
 DB 397 TTGATGAATTTACAAAGTATTTTGAATATTTTAAAGATTTTAAAGATTTTCAAG 456

QY 1427 CTATCCCTCAATCTGAGGAGCTGAGTAACACCATCATCATGATGATGAGTGGCTT 1486
 Db 457 CTATCCCTCAATCTGAGGAGCTGAGTAACACCATCATCATGATGATGAGTGGCTT 516
 QY 1487 ATGAACCTTTA 1496
 Db 517 ATGAACCTTTA 526

RESULT 15

AAS84956

ID AAS84956 standard; cDNA; 651 BP.

AC AAS84956;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #20760.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR P-PSDB; ABG20769.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

PS Claim 1; SEQ ID No 20760; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 651 BP; 152 A; 86 C; 241 G; 172 T; 0 other;

Query Match 11.6%; Score 178; DB 23; Length 651;
 Best Local Similarity 100.0%; Pred. No. 2.2e-81;

Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 661 GATCTTCTCTGATGAGGCTGACCTCCATGCCCTCTGCGAGTCTATGAGTGAAC 720
 Db 337 GATCTTCTCTGATGAGGCTGACCTCCATGCCCTCTGCGAGTCTATGAGTGAAC 396
 QY 721 AGTGAAGTGTCTGCGAGCTGTACAGAGATCGGAGACATGAGCCTCGATGCTTTTC 780
 Db 397 AGTGAAGTGTCTGCGAGCTGTACAGAGATCGGAGACATGAGCCTCGATGCTTTTC 456
 QY 781 AACATCGCCAGCTACGCCCTGCTCAAGTACATGATTTGGACATCAGGAGCTGAAGC 838
 Db 457 AACATCGCCAGCTACGCCCTGCTCAAGTACATGATTTGGACATCAGGAGCTGAAGC 514

Search completed: November 29, 2002, 00:34:55
 Job time : 468.605 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2002, 19:30:54 ; Search time 1563.2 Seconds
(without alignments)
12297.911 Million cell updates/sec

Title: US-09-963-333-7

Perfect score: 1187

Sequence: 1 gatcgccacccacccactcca.....gagggagcgccgctcgggga 1187

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estbda:*
2: em_esthum:*
3: em_estlin:*
4: em_estmni:*
5: em_estov:*
6: em_estopl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl1:*
10: gb_estl2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estlun:*
16: em_estlom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match Length | DB | ID | Description |
|------------|-------|--------------|------|----------|--------------------|
| 1 | 367.6 | 31.0 | 1061 | BQ943386 | BQ943386 AGENCOURT |
| 2 | 337.4 | 28.4 | 1060 | 13 | BM456638 AGENCOURT |
| 3 | 331.4 | 27.9 | 741 | 9 | AU118644 AU118644 |
| 4 | 329.8 | 27.8 | 739 | 9 | AU143295 AU143295 |
| 5 | 329.8 | 27.8 | 819 | 9 | AU143180 AU143180 |
| 6 | 329.4 | 27.8 | 786 | 9 | AL545940 AL545940 |

| | | | | | | |
|----|-------|------|------|----|----------|--------------------|
| 7 | 325.6 | 27.4 | 1197 | 14 | BQ072305 | BQ072305 AGENCOURT |
| 8 | 322.2 | 27.1 | 1622 | 14 | BM917794 | BM917794 AGENCOURT |
| 9 | 320.4 | 27.0 | 841 | 14 | BQ434233 | BQ434233 AGENCOURT |
| 10 | 317.4 | 26.7 | 985 | 13 | BM466216 | BM466216 AGENCOURT |
| 11 | 317.2 | 26.7 | 1018 | 14 | BM600922 | BM600922 AGENCOURT |
| 12 | 312.2 | 26.3 | 1021 | 14 | BQ685120 | BQ685120 AGENCOURT |
| 13 | 312.2 | 26.3 | 668 | 9 | AL541351 | AL541351 |
| 14 | 305.4 | 25.7 | 938 | 9 | AL551990 | AL551990 |
| 15 | 298.4 | 25.1 | 523 | 14 | BM763658 | BM763658 K-RST0045 |
| 16 | 297.4 | 25.1 | 738 | 9 | AU130205 | AU130205 |
| 17 | 295.8 | 24.9 | 711 | 9 | AU142980 | AU142980 |
| 18 | 294.2 | 24.8 | 531 | 14 | BM746027 | BM746027 K-EST0020 |
| 19 | 290.2 | 24.4 | 526 | 9 | AU128726 | AU128726 |
| 20 | 288.4 | 24.3 | 529 | 12 | BM794349 | BM794349 |
| 21 | 287.6 | 24.2 | 905 | 13 | B1597722 | B1597722 |
| 22 | 282.4 | 23.8 | 948 | 14 | BQ066205 | BQ066205 AGENCOURT |
| 23 | 281.8 | 23.7 | 935 | 9 | AL542409 | AL542409 |
| 24 | 281.2 | 23.7 | 524 | 12 | BM482405 | BM482405 |
| 25 | 280.4 | 23.6 | 1045 | 13 | BM478045 | BM478045 |
| 26 | 278.6 | 23.5 | 678 | 12 | BG337663 | BG337663 |
| 27 | 278.6 | 23.5 | 832 | 12 | BF685317 | BF685317 |
| 28 | 278.4 | 23.5 | 742 | 9 | AL548721 | AL548721 |
| 29 | 278.4 | 23.5 | 885 | 14 | BQ424746 | BQ424746 AGENCOURT |
| 30 | 277.8 | 23.4 | 1055 | 14 | BQ886916 | BQ886916 |
| 31 | 277 | 23.3 | 970 | 14 | BQ928426 | BQ928426 |
| 32 | 276.8 | 23.3 | 1020 | 9 | AL547612 | AL547612 |
| 33 | 275.8 | 23.2 | 889 | 14 | BQ420981 | BQ420981 AGENCOURT |
| 34 | 275.4 | 22.6 | 994 | 14 | BM904778 | BM904778 AGENCOURT |
| 35 | 268.2 | 22.6 | 664 | 13 | B1520187 | B1520187 |
| 36 | 267.8 | 22.6 | 406 | 14 | BM834317 | BM834317 |
| 37 | 267.8 | 22.6 | 1149 | 13 | BM546045 | BM546045 |
| 38 | 267.6 | 22.5 | 963 | 12 | BG337540 | BG337540 |
| 39 | 267.4 | 22.5 | 809 | 9 | AU117646 | AU117646 |
| 40 | 267 | 22.5 | 267 | 13 | B1668571 | B1668571 |
| 41 | 266.8 | 22.5 | 538 | 13 | BM413631 | BM413631 AGENCOURT |
| 42 | 266.6 | 22.5 | 439 | 12 | BF689564 | BF689564 |
| 43 | 265.4 | 22.4 | 672 | 12 | BG574267 | BG574267 |
| 44 | 264.6 | 22.3 | 950 | 9 | AL550812 | AL550812 |
| 45 | 264.4 | 22.3 | 688 | 9 | AL543217 | AL543217 |

ALIGNMENTS

RESULT 1
BQ943386
LOCUS BQ943386 1061 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT 8804318 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6377891
5', mRNA sequence.
ACCESSION BQ943386
VERSION BQ943386
KEYWORDS EST.
SOURCE BQ943386.1 GI:22358676
ORGANISM Homo sapiens

Hom saplens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 1061)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaps-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/ILNLI at:

<http://image.llnl.gov>

Plate: LCM562 row: 5 column: 12

High quality sequence steps: 382.

Location/Qualifiers

FEATURES

source

1. 1061

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6377891"
/clone_lib="NIH_MGC_47"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

BASE COUNT 202 a 402 c 285 g 171 t 1 others

Query Match 31.0%; Score 367.6; DB 14; Length 1061;
 Best Local Similarity 92.7%; Pred. No. 3e-36; 4; Indels 28; Gaps 1;
 Matches 408; Conservative 0; Mismatches 4;

```

QY 694 GTGCCACACCCGATGCTCTGCTTCCCTCGCGCACGCTCTTGAAGCCGCGGCGC 753
DB 1 GTGCCACACCCGATGCTCTGCTTCCCTCGCGCACGCTCTTGAAGCCGCGGCGC 60
QY 754 CGCGACCCCGCGCGAGAGAGAGCGAGCGCGGAGCGCGCGGAGAGAGCGCGG 813
DB 61 CGCGACCCCGCGCGAGAGAGAGCGAGCGCGGAGCGCGGAGAGAGAGCGCGG 120
QY 814 AAGGGGTCTTGCACCGCGCACTTGGCTGCTCCGTCGCGCGCGCACTTGGCTGC 873
DB 121 AAGGGGTCTTGCACCGCGCACTTGGCTGCTCCGTCGCGCGCGCACTTGGCTGC 152
QY 874 CTCGCTCCCGCGCGCGCACTTGGCTGCTCCGTCGCGCGCGCGCACTTGGCTGC 933
DB 153 CTCGCTCCCGCGCGCGCACTTGGCTGCTCCGTCGCGCGCGCGCACTTGGCTGC 212
QY 934 GCGGCTGGAGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 993
DB 213 GCGGCTGGAGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 272
QY 994 CCGCGCTCCCGCGCGAGCTGCACTTGGCGGAGAGAGAGAGAGAGAGAGAGAG 1053
DB 273 CCGCGCTCCCGCGCGAGCTGCACTTGGCGGAGAGAGAGAGAGAGAGAGAGAG 332
QY 1054 GCGCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1113
DB 333 GCGCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 392
QY 1114 CGCTACAGCTGAGAGGTGA 1133
DB 393 CCTACAGCTGAGAGGTGA 412

```

RESULT 2
 BM456638 1060 bp mRNA linear EST 05-FEB-2002
 LOCUS BM456638
 DEFINITION AGENCOURT 6408748 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5496315
 5', mRNA sequence.
 ACCESSION BM456638
 VERSION BM456638.1 GI:18505678
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1060)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov

FEATURES
 source
 1. 1060
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5496315"
 /clone_lib="NIH_MGC_85"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /notes="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally, oligo-dT primed.
 Average insert size 1.867 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 High quality sequence stop: 530.
 Plate: LHM2124 row: k column: 04
 Location/Qualifiers

BASE COUNT 217 a 396 c 259 g 188 t

Query Match 28.4%; Score 337.4; DB 13; Length 1060;
 Best Local Similarity 99.7%; Pred. No. 1.4e-32;
 Matches 338; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 795 CGCGGAGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 854
DB 1 CGCGGAGAGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 855 GCGCGCGCACTTGGCTGCTCTGCTCCGCGCGCGCGCACTTGGCTGCTCCGCG 914
DB 61 GCGCGCGCACTTGGCTGCTCTGCTCCGCGCGCGCGCACTTGGCTGCTCCGCG 120
QY 915 CGCGCGCGCACTTGGCTGCTCTGCTCCGCGCGCGCGCGCGCGCGCGCGCG 974
DB 121 CGCGCGCGCACTTGGCTGCTCTGCTCCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 975 ACGAGAGCGGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1034
DB 181 ACGAGAGCGGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 1035 CCAACACATCTCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1094
DB 241 CCAACACATCTCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 1095 GGTATTGCGCATGAGAGCGCGCTACAGCTTGAAGGTGA 1133
DB 301 GGTATTGCGCATGAGAGCGCGCTACAGCTTGAAGGTGA 339

```

RESULT 3
 AUI18644 741 bp mRNA linear EST 01-AUG-2002
 LOCUS AUI18644
 DEFINITION AUI18644 HEMBA1 Homo sapiens cDNA clone HEMBA1004057 5', mRNA
 sequence.
 ACCESSION AUI18644
 VERSION AUI18644.1 GI:10933771
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 741)
 AUTHORS Oca, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
 Yamamoto, V., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and
 Isogai, T.
 TITLE HRI human cDNA project
 JOURNAL Unpublished (2000)
 COMMENT Contact: Takao Isogai

Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES

source

Location/Qualifiers

1..741

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="HEMBA1004057"

/clone_id="HEMBA1"

/tissue type="whole embryo, mainly head"

/dev stage="embryo, 10 weeks"

/note="vector: pME18SF13"

BASE COUNT 153 a 227 c 218 g 140 t 3 others

ORIGIN

Query Match 27.8%; Score 331.4; DB 9; Length 741;
Best Local Similarity 99.7%; Pred. No. 9.9e-32;
Matches 332; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 801 AAAAGCGCGCGGAGAGGGTCTGCGACCGCGCCACTTGAGCTGCTCCGCGCGCG 860
Db 2 AAAAGCGCGCGGAGAGGGTCTGCGACCGCGCCACTTGAGCTGCTCCGCGCGCG 61
QY 861 CCACTTGAGCTGCTCCGCGCGCGCGCACTTGCGCTGCTCCGCGCGCGCGCG 920
Db 62 CCACTTGAGCTGCTCCGCGCGCGCGCACTTGCGCTGCTCCGCGCGCGCGCG 121
QY 921 CGCCATGCTGTGAGCGCGCTGCGAGCTGCGCGCGCGCGCTTGCCTCCCGCGCGCA 980
Db 122 CGCCATGCTGTGAGCGCGCTGCGAGCTGCGCGCGCGCGCTTGCCTCCCGCGCGCA 181
QY 981 GCGGAGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1040
Db 182 GCGGAGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 241
QY 1041 CATCTCCGCTGCGCGCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGATT 1100
Db 242 CATCTCCGCTGCGCGCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGATT 301
QY 1101 CGGATGAGAGCGCGCTACAGCTGAGAGATGA 1133
Db 302 CGGATGAGAGCGCGCTACAGCTGAGAGATGA 334

RESULT 4

AUI43295

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

HRI human cDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

FEATURES

source

Location/Qualifiers

1..739

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="Y79A1001699"

/clone_id="Y79A1"

/cell type="retinoblastoma"

/cell_line="Y79"

/note="vector: pME18SF13"

BASE COUNT 155 a 223 c 217 g 140 t 4 others

ORIGIN

Query Match 27.8%; Score 329.8; DB 9; Length 739;
Best Local Similarity 99.4%; Pred. No. 1.6e-31;
Matches 331; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 801 AAAAGCGCGCGGAGAGGGTCTGCGACCGCGCCACTTGAGCTGCTCCGCGCGCG 860
Db 2 AAAAGCGCGCGGAGAGGGTCTGCGACCGCGCCACTTGAGCTGCTCCGCGCGCG 61
QY 861 CCACTTGAGCTGCTCCGCGCGCGCGCACTTGCGCTGCTCCGCGCGCGCGCG 920
Db 62 CCACTTGAGCTGCTCCGCGCGCGCGCACTTGCGCTGCTCCGCGCGCGCGCG 121
QY 921 CGCCATGCTGTGAGCGCGCTGCGAGCTGCGCGCGCGCGCTTGCCTCCCGCGCGCA 980
Db 122 CGCCATGCTGTGAGCGCGCTGCGAGCTGCGCGCGCGCGCTTGCCTCCCGCGCGCA 181
QY 981 GCGGAGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1040
Db 182 GCGGAGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 241
QY 1041 CATCTCCGCTGCGCGCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGATT 1100
Db 242 CATCTCCGCTGCGCGCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGATT 301
QY 1101 CGGATGAGAGCGCGCTACAGCTGAGAGATGA 1133
Db 302 CGGATGAGAGCGCGCTACAGCTGAGAGATGA 334

RESULT 5

AUI43180

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

HRI human cDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES

Location/Qualifiers

1..819

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="Y79AA1001513"

/clone_lib="Y79AA1"

/cell_type="retinoblastoma"

/note="Vector: pME18SFL3"

/note="Vector: pME18SFL3"

BASE COUNT 169 a 243 c 239 g 164 t 4 others

ORIGIN

Query Match 27.8%; Score 329.8; DB 9; Length 819;
Best Local Similarity 99.4%; Pred. No. 1.4e-31;
Matches 331; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 801 AAAAGGCGCGGAGAGGGGTCTGCGACCGGCACTTGGCTGCTCCGCGCG 860
DB 2 AAAAGGCGCGGAGAGGGGTCTGCGACCGGCACTTGGCTGCTCCGCGCG 61
QY 861 CCACTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
DB 62 CCACTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
QY 921 CGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 980
DB 122 CGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 101
QY 981 GCGGAGCGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1040
DB 182 GCGGAGCGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241
QY 1041 CATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1100
DB 242 CATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
QY 1101 CGGCATGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1133
DB 302 CGGCATGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 334

RESULT 6
AL545940 786 bp mRNA linear EST 16-FEB-2001
LOCUS AL545940 LTR N1006 PL2 Homo sapiens cDNA clone CS0D1023YK20 5

DEFINITION
prime, mRNA sequence.

ACCESSION
AL545940
VERSION
AL545940.1 GI:12878592
KEYWORDS
EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE
1 (bases 1 to 786)
AUTHORS
Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished (2001)

COMMENT

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

Location/Qualifiers

1..786

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0D1023YK20"

/clone_lib="LTR N1006 PL2"

/tissue_type="Placenta"

/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA"

was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

life Technologies. Contact : Peng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : fliang@life-tech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 159 a 233 c 233 g 152 t 9 others

ORIGIN

Query Match 27.8%; Score 329.4; DB 9; Length 786;
Best Local Similarity 99.1%; Pred. No. 1.6e-31;
Matches 330; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 801 AAAAGGCGCGGAGAGGGGTCTGCGACCGGCACTTGGCTGCTCCGCGCG 860
DB 1 AAAAGGCGCGGAGAGGGGTCTGCGACCGGCACTTGGCTGCTCCGCGCG 60
QY 861 CCACTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
DB 61 CCACTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 921 CGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 980
DB 121 CGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 981 GCGGAGCGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1040
DB 181 GCGGAGCGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 1041 CATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1100
DB 241 CATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 1101 CGGCATGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1133
DB 301 CGGCATGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 333

RESULT 7
B0072305 1197 bp mRNA linear EST 02-APR-2002
LOCUS B0072305 AGENCOURT 6838866 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5761428

DEFINITION
5', mRNA sequence.

ACCESSION
B0072305
VERSION
B0072305.1 GI:19901351
KEYWORDS
EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE
1 (bases 1 to 1197)
AUTHORS
NIH-MGC http://mgi.mol.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cga@bbsr@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLM at:
http://image.llnl.gov

Plate: IMAGE2810 row: a column: 13
High quality sequence stop: 200.

FEATURES

Location/Qualifiers

1..1197

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5761428"


```

/clone.lib="NIH_MGC_122"
/lab_host="DH10B"
/notes="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site 1: NotI; Site 2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH MGC Library."

BASE COUNT      274 a      310 c      493 g      120 t
ORIGIN

Query Match      27.4%; Score 325.6; DB 14; Length 1197;
Best Local Similarity 98.8%; Pred. No. 3.3e-31;
Matches 328; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 802 AAAGGCGCGGAAAGGGTCTGTCACCCGCGCATTGAGCTCGCTCCGCGCGCGC 861
Db 1 AAAGCGCGGAAAGGGTCTGTCACCCGCGCATTGAGCTCGCTCCGCGCGCGC 60
QY 862 CACTTGGCGCTGCTCCGCTCCGCGCGCATTGAGCTCGCTCCGCGCGCGC 921
Db 61 CACTTGGCGCTGCTCCGCTCCGCGCGCATTGAGCTCGCTCCGCGCGCGC 120
QY 922 GCGATGCTGTGGCGCGCTGCGAGACTGCGCGCGCGCTTGGCCCGCGCGCAGAG 981
Db 121 GCGATGCTGTGGCGCGCTGCGAGACTGCGCGCGCGCTTGGCCCGCGCGCAGAG 180
QY 982 CGGAGCGCGGAGCGCGCTGCGCGCAGAGACTGCGAGTACTTGGGCGAGATCCACAC 1041
Db 181 CGGAGCGCGGAGCGCGCTGCGCGCAGAGACTGCGAGTACTTGGGCGAGATCCACAC 240
QY 1042 ATCTCTCGCTGCGCGCTGCGAGAGAGACGACCGCGAGCGCGCTGTGTGGATTTC 1101
Db 241 ATCTCTCGCTGCGCGCTGCGAGAGAGAGACCGCGAGCGCGCTGTGTGGATTTC 300
QY 1102 GCGATGCGAGCGCGCTTACAGCTTGAAGGTGA 1133
Db 301 GCGATGCGAGCGCGCTTACAGCTTGAAGGTGA 332

RESULT 8
BM917794      1622 bp      mRNA      linear      EST 12-MAR-2002
LOCUS      AGENCOURT 6614511 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5485129
DEFINITION      5', mRNA sequence.
ACCESSION      BM917794
VERSION      BM917794.1 GI:19368173
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      1 (bases 1 to 1622)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DMS/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1CM2015 row: i column: 02
High quality sequence stop: 431.
Location/Qualifiers
1..1622

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5485129"
/clone_id="NIH_MGC_106"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: blood; Vector: pOTB1; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCGAGG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."

BASE COUNT      430 a      610 c      271 g      289 t      22 others
ORIGIN

Query Match      27.1%; Score 322.2; DB 14; Length 1622;
Best Local Similarity 99.1%; Pred. No. 6.5e-31;
Matches 324; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 807 CGGCGGAAAGGGTCTCTGCGACCGCGCATTGAGCTGTCCGCGCGCGCACTT 866
Db 1 CGGCGGAAAGGGTCTCTGCGACCGCGCATTGAGCTGTCCGCGCGCGCACTT 60
QY 867 GCGCTGCTGCTGCTCCGCGCGCGCATTGAGCTGTCCGCGCGCGCGCGCAT 926
Db 61 GCGCTGCTGCTGCTCCGCGCGCGCATTGAGCTGTCCGCGCGCGCGCGCAT 120
QY 927 GCGTGTGCGCGGCTGCGAGCTGCGCGCGCGGCGGCTTGGCCCGCGCGCAGAGCGGGA 986
Db 121 GCGTGTGCGCGGCTGCGAGCTGCGCGCGCGGCGGCTTGGCCCGCGCGCAGAGCGGGA 180
QY 987 CGCGGAGCGCGCTGCGCGCAGAGAGCTGAGTACTTGGGCGAGATCCACATCTT 1046
Db 181 CGCGGAGCGCGCTGCGCGCAGAGAGCTGAGTACTTGGGCGAGATCCACATCTT 240
QY 1047 CCGCTGTGGCGTGTGAGAAAGACGACCGCGAGCGCGCAGCTGTGTGGATTTC 1106
Db 241 CCGCTGTGGCGTGTGAGAAAGACGACCGCGAGCGCGCAGCTGTGTGGATTTC 300
QY 1107 GCGAGCGCGCTTACAGCTTGAAGGTGA 1133
Db 301 GCGAGCGCGCTTACAGCTTGAAGGTGA 327

RESULT 9
BQ434233      841 bp      mRNA      linear      EST 24-MAY-2002
LOCUS      AGENCOURT 7896761 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6159722
DEFINITION      5', mRNA sequence.
ACCESSION      BQ434233
VERSION      BQ434233.1 GI:21173309
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      1 (bases 1 to 841)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM13509 row: e column: 03
High quality sequence stop: 639.

```

FEATURES

source

Location/Qualifiers

1. 841
 /organism="Homo sapiens"
 /db xref="taxon:9606"
 /clone="IMAGE:6159722"
 /tissue_lib="NIH MGC 72"
 /tissue_type="melanotic melanoma"
 /lab host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: PCMV-SPOrt6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

BASE COUNT 176 a 263 c 241 g 161 t
 ORIGIN

Query Match 27.0%; Score 320.4; DB 14; Length 841;
 Best Local Similarity 99.7%; Pred. No. 1.9e-30;
 Matches 321; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 812 GGAAGGGGTCCTCCACCGGCGCACTTGACCTGCTCCGCGCGCCCACTTGAGCT 871
 DB 1 GGAAGGGGTCCTCCACCGGCGCACTTGACCTGCTCCGCGCGCCCACTTGAGCT 60
 QY 872 GCTCCGTCCTCCGCGCGCACTTGACCTGCTCCGCGCGCCCACTTGAGCT 931
 DB 61 GCTCCGTCCTCCGCGCGCACTTGACCTGCTCCGCGCGCCCACTTGAGCT 120
 QY 932 TGGCGGCTCGAGACTGCGCGCGCGCTTGCCTCCGCGCGCAAGAGCGGACGCG 991
 DB 121 TGGCGGCTCGAGACTGCGCGCGCGCTTGCCTCCGCGCGCAAGAGCGGACGCG 180
 QY 992 AGCCGCGTCCGCGCGCAAGAGACTGAGTATGAGGCGAGATCCAACTTCCTCCCT 1051
 DB 181 AGCCGCGTCCGCGCGCAAGAGACTGAGTATGAGGCGAGATCCAACTTCCTCCCT 240
 QY 1052 GGGGCGTGAAG 1111
 DB 241 GGGGCGTGAAG 300
 QY 1112 CGCGCTACAGCCTGAGAGGTGA 1133
 DB 301 CGCGCTACAGCCTGAGAGATGA 322

RESULT 10

BM66216

LOCUS BM66216 985 bp mRNA linear EST 05-FEB-2002

DEFINITION AGENCOURT_6456956 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5576962

5', mRNA sequence.

ACCESSION BM66216

VERSION BM66216.1 GI:18515258

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 985)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LHAM12329 row: k column: 11

High quality sequence stop: 522.

Location/Qualifiers

1. 985

source

/organism="Homo sapiens"

/db xref="taxon:9606"

/clone="IMAGE:5576962"

/tissue_lib="NIH MGC 92"

/tissue_type="embryonal carcinoma, cell line"

/lab host="DH10B (phage-resistant)"

/note="Organ: testis; Vector: PCMV-SPOrt6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally. Oligo-dt primed.

Average insert size 2.5 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH MGC library."

BASE COUNT 194 a 323 c 264 g 204 t
 ORIGIN

Query Match 26.7%; Score 317.4; DB 13; Length 985;
 Best Local Similarity 98.2%; Pred. No. 3.9e-30;
 Matches 321; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 807 CGCGGGAAGGGGTCCTCCACCGGCGCACTTGACCTGCTCCGCGCGCCCACTTGAGCT 866
 DB 4 CGCGGTCCTCCGCGCGCACTTGACCTGCTCCGCGCGCCCACTTGAGCT 63
 QY 867 GGCCTGCTCCGCGCGCACTTGACCTGCTCCGCGCGCCCACTTGAGCT 926
 DB 64 GGCCTGCTCCGCGCGCACTTGACCTGCTCCGCGCGCCCACTTGAGCT 123
 QY 927 GCTGTGCTCGAGACTGCGCGCGCGCTTGCCTCCGCGCGCAAGAGCGGACGCG 986
 DB 124 GCTGTGCTCGAGACTGCGCGCGCGCTTGCCTCCGCGCGCAAGAGCGGACGCG 183
 QY 987 CGCGGAGCGGTCCTCCGCGCGCAAGAGACTGAGTATGAGGCGAGATCCAACTTCCT 1046
 DB 184 CGCGGAGCGGTCCTCCGCGCGCAAGAGACTGAGTATGAGGCGAGATCCAACTTCCT 243
 QY 1047 CGCTGCGGCTCGAG 1106
 DB 244 CGCTGCGGCTCGAG 303
 QY 1107 GCAAGGCGGCTACAGCCTGAGAGGTGA 1133
 DB 304 GCAAGGCGGCTACAGCCTGAGAGATGA 330

RESULT 11

BM800322

LOCUS BM800322 1018 bp mRNA linear EST 05-MAR-2002

DEFINITION AGENCOURT_6420080 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5584389

5', mRNA sequence.

ACCESSION BM800322

VERSION BM800322.1 GI:19117145

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1018)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LHAM12348 row: p column: 22

High quality sequence stop: 583.

Location/Qualifiers

1. 1018

/organism="Homo sapiens"

source

/db_xref="taxon:9606"
 /clone_image="IMAGE:5584389"
 /clone_lib="NIH_MGC_92"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
 BASE COUNT 217 a 322 c 273 g 206 t
 ORIGIN

Query Match 26.7%; Score 317.2; DB 14; Length 1018;
 Best Local Similarity 99.1%; Pred. No. 4e-30; Mismatches 3; Indels 0; Gaps 0;
 Matches 319; Conservative 0;

QY 812 GGAAGGGGCTCTGCCACCGGCACTTGCGCTCGCTCGGCGCGCGCGCACTTGAGCT 871
 Db 2 GGAAGGGGCTCTGCCACCGGCACTTGCGCTCGCTCGGCGCGCGCGCACTTGAGCT 61
 QY 872 GCGTCCTCCGCGCGCGCGCACTTGCGCTCGCTCGGCGCGCGCGCACTTGAGCT 931
 Db 62 GCGTCCTCCGCGCGCGCGCACTTGCGCTCGCTCGGCGCGCGCGCACTTGAGCT 121
 QY 932 TGCGCGGCTCGGAGCTCGCGCGCGCGCTTGCGCGCGCGCGCACTTGAGCT 991
 Db 122 TGCGCGGCTCGGAGCTCGCGCGCGCGCTTGCGCGCGCGCGCACTTGAGCT 181
 QY 992 AGCGCGGCTCGCGCGCGCGCACTTGCGCTTGCGCGCGCGCGCACTTGAGCT 1051
 Db 182 AGCGCGGCTCGCGCGCGCGCACTTGCGCTTGCGCGCGCGCGCACTTGAGCT 241
 QY 1052 GCGGCGTGAAG 1111
 Db 242 GCGGCGTGAAG 301
 QY 1112 CGCGCTACAGCTTGAGAGCTGA 1133
 Db 302 CGCGCTACAGCTTGAGAGATGA 323

RESULT 12 B0685120 1021 bp mRNA linear EST 15-JUL-2002
 LOCUS B0685120
 DEFINITION AGENCOURT 8344472 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6251301
 5', mRNA sequence.
 ACCESSION B0685120
 VERSION B0685120.1 GI:21810436
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1021)
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: L10M2395 row: 1 column: 22
 High quality sequence start: 12
 High quality sequence stop: 409.
 Location/Qualifiers

FEATURES
 1..1021
 source /organism="Homo sapiens"

/db_xref="taxon:9606"
 /clone_image="IMAGE:6251301"
 /clone_lib="NIH_MGC_110"
 /tissue_type="ductal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: pancreas; Vector: pORF7; Site 1: XhoI; Site 2: SacI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
 BASE COUNT 202 a 335 c 289 g 194 t 1 others
 ORIGIN

Query Match 26.3%; Score 312.2; DB 14; Length 1021;
 Best Local Similarity 99.1%; Pred. No. 1.6e-29; Mismatches 3; Indels 0; Gaps 0;
 Matches 314; Conservative 0;

QY 817 GGGTCTGCGACCGGCGCACTTGAGCTCGCTCGGCGCGCGCGCACTTGAGCT 876
 Db 7 GGGTCTGCGACCGGCGCACTTGAGCTCGCTCGGCGCGCGCGCACTTGAGCT 66
 QY 877 CGTCCCGCGCGCGCACTTGCGCTCGCTCGGCGCGCGCGCACTTGAGCT 936
 Db 67 CGTCCCGCGCGCGCACTTGCGCTCGCTCGGCGCGCGCGCACTTGAGCT 126
 QY 937 GGGTGGAGACTGCGCGCGCGCTTGCGCGCGCGCGCACTTGAGCT 996
 Db 127 GGGTGGAGACTGCGCGCGCGCTTGCGCGCGCGCGCACTTGAGCT 186
 QY 997 CGTCCCGCGCACTTGCGAGCTGCACTTGCGGCGAGATCAACATCTCTCGCTG 1056
 Db 187 CGTCCCGCGCACTTGCGAGCTGCACTTGCGGCGAGATCAACATCTCTCGCTG 246
 QY 1057 GTGAG 1116
 Db 247 GTGAG 306
 QY 1117 TACAGCTTGAGAGGTTGA 1133
 Db 307 TACAGCTTGAGAGATGA 323

RESULT 13 AL541351 668 bp mRNA linear EST 16-FEB-2001
 LOCUS AL541351
 DEFINITION AL541351 LTI FL002_P11 Homo sapiens cDNA clone CSDBE006YA08 5 prime
 , mRNA sequence.
 ACCESSION AL541351
 VERSION AL541351.1 GI:12872339
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 668)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segreg@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 1..668
 source /organism="Homo sapiens"

/db_xref="taxon:9606"
 /clone_image="CSDBE006YA08"
 /clone_lib="LTI FL002_P11"
 /lab_host="DH10B"
 /note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand

cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com>

BASE COUNT 139 a 208 c 192 g 126 t 3 others

ORIGIN

Query Match 26.3%; Score 312; DB 9; Length 668;
Best Local Similarity 99.4%; Pred. No. 2,4e-29;
Matches 312; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 820 TCCTGCGACCGCGCACTTGGCTGCTCCGTCGCGCGGCACTTGGCTGCTCCGCT 879
Db 1 TCTGCGACCGCGCACTTGGCTGCTCCGTCGCGCGGCACTTGGCTGCTCCGCT 60
QY 880 CCGCGCGCGCACTTGGCTGCTCCGTCGCGCGGCACTTGGCTGCTCCGCT 939
Db 61 CCGCGCGCGCACTTGGCTGCTCCGTCGCGCGGCACTTGGCTGCTCCGCT 120
QY 940 TCGGAGCTGCG 999
Db 121 TCGGAGCTGCG 180
QY 1000 CCG 1059
Db 181 CCG 240
QY 1060 AGGAGAGACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1119
Db 241 AGGAGAGACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 1120 AGCTTGAAGGTGA 1133
Db 301 AGCTTGAAGGTGA 314

RESULT 14 938 bp mRNA linear EST 16-FEB-2001
AL551990 LOCUS
DEFINITION AL551990 LTI NFL006 PL2 Homo sapiens cDNA clone CS01060Y04 5
prime mRNA sequence.

ACCESSION AL551990
VERSION AL551990.1 GI:12890459
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 938)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1. 938
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS01060Y04"
/clone_1ib="LTI_NFL006 PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by

Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com>

BASE COUNT 202 a 288 c 259 g 188 t 1 others

ORIGIN

Query Match 25.7%; Score 305.4; DB 9; Length 938;
Best Local Similarity 99.7%; Pred. No. 1.1e-28;
Matches 306; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 827 ACCGCGCACCTTGGCTGCTCCGTCGCGCGGCACTTGGCTGCTCCGCT 886
Db 1 ACCGCGCACCTTGGCTGCTCCGTCGCGCGGCACTTGGCTGCTCCGCT 60
QY 887 CGCCACTTGGCTGCTCCGTCGCTCCGTCGCGCGGCACTTGGCTGCTCCGCT 946
Db 61 CGCCACTTGGCTGCTCCGTCGCTCCGTCGCGCGGCACTTGGCTGCTCCGCT 120
QY 947 TCG 1006
Db 121 TCG 180
QY 1007 AGGAGAGACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1066
Db 181 AGGAGAGACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
QY 1067 AGGAGAGACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1126
Db 241 AGGAGAGACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 1127 GAGGTGA 1133
Db 301 GAGGTGA 307

RESULT 15 523 bp mRNA linear EST 04-MAR-2002
BM763658 LOCUS
DEFINITION K-EST0045041 S13XMS5 Homo sapiens cDNA clone S13XMS5-14-F09 5',
mRNA sequence.

ACCESSION BM763658
VERSION BM763658.1 GI:19093273
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 523)
AUTHORS Kim, M.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.krrib.re.kr
Plate: 14 row: F Column: 09
High quality sequence stop: 523.

FEATURES

source

1. 523
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S13XMS5-14-F09"
/clone_1ib="S13XMS5"
/tissue_type="myeloma"
/cell_line="KMS-5"
/lab_host="Top10F"

/note="Vector: pCNS; Site_1: EcoRI; Site_2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT 89 a 173 c 163 g 98 t
ORIGIN

Query Match 25.1%; Score 298.4; DB 14; Length 523;
Best Local Similarity 99.7%; Pred. No. 1.4e-27;
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 834 CACTTGGCTGCTCCGTCCTCCGCGCGCCACTTGGCTGCTCCGTCCTCCGCGCGCCACT 893
    |||
DB 2 CACTTGGCTGCTCCGTCCTCCGCGCGCCACTTGGCTGCTCCGTCCTCCGCGCGCCACT 61

QY 894 TCGCTGCTCCGTCCTCCGCGCGCGCCACTTGGCTGCTCCGTCCTCCGCGCGCCACT 953
    |||
DB 62 TCGCTGCTCCGTCCTCCGCGCGCGCCACTTGGCTGCTCCGTCCTCCGCGCGCCACT 121

QY 954 CCGGCTGCTCCGTCCTCCGCGCGCGCCACTTGGCTGCTCCGTCCTCCGCGCGCCACT 1013
    |||
DB 122 CCGGCTGCTCCGTCCTCCGCGCGCGCCACTTGGCTGCTCCGTCCTCCGCGCGCCACT 181

QY 1014 GCTGAGTACCTGGGAGATCAACATCTCTGCGGCGTCAAGAGAGACGACCG 1073
    |||
DB 182 GCTGAGTACCTGGGAGATCAACATCTCTGCGGCGTCAAGAGAGACGACCG 241

QY 1074 CACGGGACCGGACCGCTGCGGATTCGCGATGAGCGCGCTACAGCTGAGAGGTGA 1133
    |||
DB 242 CACGGGACCGGACCGCTGCGGATTCGCGATGAGCGCGCTACAGCTGAGAGATGA 301
```

Search completed: November 28, 2002, 22:40:23
Job time : 1582.2 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2002, 19:48:24 ; Search time 92.4142 Seconds

(without alignments)
8693.659 Million cell updates/sec

Title: US-09-963-333-7

Perfect score: 1187
Sequence: 1 gatcgccacatcgactcca.....gagggagcgcgctgggga 1187Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 344316 segs, 338423730 residues

Total number of hits satisfying chosen parameters: 688632

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA New:*
1: /cgn2_6/ptodata/2/pna/US06_NMW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NMW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NMW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US09_NMW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US10_NMW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US60_NMW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US60_NMW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------------------|--------------------|
| 1 | 1178 | 99.2 | 1187 | 5 | US-09-658-659B-7 | Sequence 7, Appli |
| 2 | 1167 | 98.3 | 18597 | 5 | US-09-658-659B-8 | Sequence 8, Appli |
| 3 | 770.6 | 64.9 | 1186 | 6 | US-10-257-166-172 | Sequence 172, Appl |
| 4 | 710 | 59.8 | 1186 | 6 | US-10-257-166-171 | Sequence 171, Appl |
| 5 | 301 | 25.4 | 1942 | 5 | US-09-724-676-34523 | Sequence 34523, A |
| 6 | 301 | 25.4 | 1942 | 5 | US-09-724-676A-34523 | Sequence 34523, A |
| 7 | 301 | 25.4 | 2042 | 5 | US-09-724-676-34515 | Sequence 34515, A |
| 8 | 301 | 25.4 | 2042 | 5 | US-09-724-676A-34515 | Sequence 34515, A |
| 9 | 301 | 25.4 | 2391 | 5 | US-09-724-676A-34597 | Sequence 34597, A |
| 10 | 301 | 25.4 | 2391 | 5 | US-09-724-676A-34597 | Sequence 34597, A |
| 11 | 301 | 25.4 | 2411 | 5 | US-09-724-676-34535 | Sequence 34535, A |
| 12 | 301 | 25.4 | 2411 | 5 | US-09-724-676A-34535 | Sequence 34535, A |
| 13 | 301 | 25.4 | 2491 | 5 | US-09-724-676-34589 | Sequence 34589, A |
| 14 | 301 | 25.4 | 2491 | 5 | US-09-724-676A-34589 | Sequence 34589, A |
| 15 | 301 | 25.4 | 2527 | 5 | US-09-724-676-34548 | Sequence 34548, A |
| 16 | 301 | 25.4 | 2527 | 5 | US-09-724-676A-34548 | Sequence 34548, A |
| 17 | 301 | 25.4 | 3023 | 5 | US-09-724-676-34565 | Sequence 34565, A |
| 18 | 301 | 25.4 | 3023 | 5 | US-09-724-676A-34565 | Sequence 34565, A |
| 19 | 301 | 25.4 | 3123 | 5 | US-09-724-676-34557 | Sequence 34557, A |
| 20 | 301 | 25.4 | 3123 | 5 | US-09-724-676A-34557 | Sequence 34557, A |
| 21 | 301 | 25.4 | 3139 | 5 | US-09-724-676-34581 | Sequence 34581, A |
| 22 | 301 | 25.4 | 3139 | 5 | US-09-724-676A-34581 | Sequence 34581, A |
| 23 | 301 | 25.4 | 3239 | 5 | US-09-724-676-34573 | Sequence 34573, A |
| 24 | 301 | 25.4 | 3239 | 5 | US-09-724-676A-34573 | Sequence 34573, A |
| 25 | 299.4 | 25.2 | 1536 | 5 | US-09-658-659B-6 | Sequence 6, Appli |
| 26 | 288.8 | 24.3 | 3298 | 6 | US-10-240-965-200 | Sequence 200, Appl |

| | | | | | | |
|----|-------|------|--------|---|----------------------|---------------------|
| 27 | 204 | 17.2 | 304905 | 1 | PCT-US02-32700-1 | Sequence 1, Appl |
| 28 | 204 | 17.2 | 304905 | 6 | US-10-271-416-1 | Sequence 1, Appl |
| 29 | 199 | 16.8 | 135259 | 6 | US-10-240-425-1585 | Sequence 1585, Ap |
| 30 | 198.4 | 16.7 | 207433 | 6 | US-10-277-216-5 | Sequence 5, Appl |
| 31 | 193.4 | 16.3 | 162025 | 6 | US-10-272-665-35 | Sequence 35, Appl |
| 32 | 193.4 | 16.3 | 162025 | 6 | US-10-272-665-36 | Sequence 36, Appl |
| 33 | 193.4 | 16.3 | 162025 | 6 | US-10-273-321-35 | Sequence 35, Appl |
| 34 | 193.4 | 16.3 | 162025 | 6 | US-10-273-321-36 | Sequence 36, Appl |
| 35 | 193.4 | 16.3 | 162025 | 6 | US-10-272-756-35 | Sequence 35, Appl |
| 36 | 193.4 | 16.3 | 162025 | 6 | US-10-272-756-36 | Sequence 36, Appl |
| 37 | 193.4 | 16.3 | 162025 | 6 | US-10-273-228-35 | Sequence 35, Appl |
| 38 | 193.4 | 16.3 | 162025 | 6 | US-10-273-228-36 | Sequence 36, Appl |
| 39 | 192.6 | 16.2 | 202100 | 1 | PCT-US02-34679-484 | Sequence 36, Appl |
| 40 | 192.6 | 16.2 | 202100 | 6 | US-10-282-174-484 | Sequence 484, Appl |
| 41 | 191.2 | 16.1 | 36534 | 6 | US-10-240-425-1461 | Sequence 1461, Ap |
| 42 | 191.2 | 16.1 | 145831 | 6 | US-10-240-425-1461 | Sequence 1461, Ap |
| 43 | 191 | 16.1 | 10067 | 6 | US-10-240-425-1470 | Sequence 1470, Appl |
| 44 | 189.2 | 15.9 | 396 | 5 | US-09-513-999C-28575 | Sequence 28575, A |
| 45 | 189.2 | 15.9 | 201143 | 6 | US-10-240-425-1099 | Sequence 1099, Ap |

ALIGNMENTS

RESULT 1
US-09-658-659B-7
Sequence 7, Application US/09658659B
GENERAL INFORMATION:
APPLICANT: Stanton, Jr., Vincent P.
TITLE OF INVENTION: GENE SEQUENCE VARIANCE IN GENES RELATED
TO FOLATE METABOLISM HAVING UTILITY IN DETERMINING THE
TITLE OF INVENTION: TREATMENT OF DISEASE
FILE REFERENCE: 11926-015001
CURRENT APPLICATION NUMBER: US/09/658,659B
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 09/596,033
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 09/357,743
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 09/357,024
PRIOR FILING DATE: 1999-07-19
PRIOR APPLICATION NUMBER: 60/093,484
PRIOR FILING DATE: 1998-07-20
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 1187
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 276_321, 534, 656
OTHER INFORMATION: n = c or t
FEATURE:
NAME/KEY: misc.feature
LOCATION: 452, 640
OTHER INFORMATION: n = a or g
FEATURE:
NAME/KEY: misc.feature
LOCATION: 492, 625
OTHER INFORMATION: n = c or a
OTHER INFORMATION: nucleotide in position 458 is c, or absent
US-09-658-659B-7
Query Match 99.2%; Score 1178; DB 5; Length 1187;
Best Local Similarity 100.0%; Pred. No. 4.8e-151;
Matches 1187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GATCGGCACGACGACCTGAGGAGGAGCGGAGCTGCTGCTCAAAAAA 60
|||||

Dh 1 GATCGCGCACTGACCTCCAGCTGGGTGAGAGCAGACTGTCTCTCAAAAAA 60
Qy AAAAGACCGCCAGGCTCAAAACAAAAACCTCGAAAAAGCCCTGGGGCTTTT 120
Dh 61 AAAAGACCGCCAGGCTCAAAACAAAAACCTCGAAAAAGCCCTGGGGCTTTT 120
Qy 121 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 180
Dh 121 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 180
Qy 181 CGGACTTGGCTCACTGAACTCTGCTCCAGGTTCAACATTTCTTCTGCTAGCC 240
Dh 181 CGGACTTGGCTCACTGAACTCTGCTCCAGGTTCAACATTTCTTCTGCTAGCC 240
Qy 241 TCCCAAGTACCAACGCGCCAGCTAATTTTGTATTTTATGAGACGCGGGTTTCA 300
Dh 241 TCCCAAGTACCAACGCGCCAGCTAATTTTGTATTTTATGAGACGCGGGTTTCA 300
Qy 301 CATGTTGTCAGAGCTGCTTNGAACTCTGACCTCAAGTATCAACCCGCTCGGCC 360
Dh 301 CATGTTGTCAGAGCTGCTTNGAACTCTGACCTCAAGTATCAACCCGCTCGGCC 360
Qy 361 CAAAGTACTAGATTACAGGCGTGAAGCAGCCGCTCAGCGCCCTGGGGTTTATCA 420
Dh 361 CAAAGTACTAGATTACAGGCGTGAAGCAGCCGCTCAGCGCCCTGGGGTTTATCA 420
Qy 421 AGTAAAAAGCTGATTAACCTGCTTNGCTTNGCTTNGCTTNGCTTNGCTTNG 480
Dh 421 AGTAAAAAGCTGATTAACCTGCTTNGCTTNGCTTNGCTTNGCTTNGCTTNG 480
Qy 481 AATGGAATCCTTATGTTGTAAGAAACAGATCTCAACAGCACTTTTGTGACAG 540
Dh 481 AATGGAATCCTTATGTTGTAAGAAACAGATCTCAACAGCACTTTTGTGACAG 540
Qy 541 ACCGAGAGAAAAAGCTGGAACCTGCTGCTTGAAGAAAGCGCGTGCACAGACG 600
Dh 541 ACCGAGAGAAAAAGCTGGAACCTGCTGCTTGAAGAAAGCGCGTGCACAGACG 600
Qy 601 TTCCCAAGGCGGAGCTCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 660
Dh 601 TTCCCAAGGCGGAGCTCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 660
Qy 661 AAGACTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Dh 661 AAGACTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy 721 CCGCTGCGGAGCTCTGAGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
Dh 721 CCGCTGCGGAGCTCTGAGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
Qy 781 GAGCGGAGGAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Dh 781 GAGCGGAGGAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Qy 841 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Dh 841 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Qy 901 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Dh 901 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Qy 961 TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Dh 961 TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Qy 1021 TACTGCGGAGATCAACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Dh 1021 TACTGCGGAGATCAACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Qy 1081 ACCGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Dh 1081 ACCGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140

Qy 1141 GATCGCGCACTGACCTCCAGCTGGGTGAGAGCAGACTGTCTCTCAAAAAA 1187
Dh 1141 GATCGCGCACTGACCTCCAGCTGGGTGAGAGCAGACTGTCTCTCAAAAAA 1187

RESULT 2
US-09-658-659B-8
; Sequence 8, Application US/09658659B
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: GENE SEQUENCE VARIANCE IN GENES RELATED
; TITLE OF INVENTION: TO FOLATE METABOLISM HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015001
; CURRENT APPLICATION NUMBER: US/09/658,659B
; CURRENT FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 18597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 701..13751
; OTHER INFORMATION: n = c or a
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 716, 1293, 2401, 2429, 2618, 3083, 3125, 3635, 4256, 4898,
; LOCATION: 5062, 5167, 11069, 13298, 14479, 14730, 14796, 15344, 15450,
; LOCATION: 15503, 15590, 15840, 16149
; OTHER INFORMATION: n = a or g
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 732, 1379, 1590, 2488, 3212, 5006, 11238, 11422, 11686,
; LOCATION: 12598, 13171, 13645, 13782, 13806, 13813, 14586, 14788,
; LOCATION: 15042, 15546, 15770
; OTHER INFORMATION: n = c or t
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1322, 1688
; OTHER INFORMATION: n = c or g
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 2594, 11293, 16199, 16203
; OTHER INFORMATION: n = g or t
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 3619
; OTHER INFORMATION: n = a or t
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 14547
; OTHER INFORMATION: nucleotide in position 14547 is t, or absent
US-09-658-659B-8

Query Match 98.3%; Score 1167; DB 5; Length 18597;
Best local Similarity 99.5%; Pred. No. 1,1e-145;
Matches 1181; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1 GATCGCGCACTGACCTCCAGCTGGGTGAGAGCAGACTGTCTCTCAAAAAA 60
Dh 78 GATCGCGCACTGACCTCCAGCTGGGTGAGAGCAGACTGTCTCTCAAAAAA 137

QY 61 AAAAGACCGCCAGGCTCAAAACAAAAAAGCTTGGAAAAAGCTTGGGCTTTT 120
Db 138 AAAAGACCGCCAGGCTCAAAACAAAAAAGCTTGGAAAAAGCTTGGGCTTTT 197
QY 121 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 180
Db 198 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 257
QY 181 CGAGTCTGGCTCACTGCAACCTTGGCTCCAGGTTCAAGCAATCTTCTGCTGAGCC 240
Db 258 CGAGTCTGGCTCACTGCAACCTTGGCTCCAGGTTCAAGCAATCTTCTGCTGAGCC 317
QY 241 TCCCAAGTAGCCACACGCCAGCTAAATTTTGTATTTTGTATAGAGACGGGGTTTCA 300
Db 318 TCCCAAGTAGCCACACGCCAGCTAAATTTTGTATTTTGTATAGAGACGGGGTTTCA 377
QY 301 CATGTTGTCCAGGCTGTCTTGAACCTCTGACTCAAGGATGATCCAGCTGCGCCGCC 360
Db 378 CATGTTGTCCAGGCTGTCTTGAACCTCTGACTCAAGGATGATCCAGCTGCGCCGCC 437
QY 361 CAAAGTACTAGATTTACAGGGGTGAGGCAACGGCTCAAGCCCTGGGGTTTATATCA 420
Db 438 CAAAGTACTAGATTTACAGGGGTGAGGCAACGGCTCAAGCCCTGGGGTTTATATCA 497
QY 421 AATAGAAAAGCTGATTAATACCTTGGCTTGGTGGCTTGGTGAAGAGAAATG 480
Db 498 AATAGAAAAGCTGATTAATACCTTGGCTTGGTGGCTTGGTGAAGAGAAATG 556
QY 481 AATAGCAATCNCCTATTAGTTAGTGAAGAAACAGATCTCAACAGAGTTTGTGACAG 540
Db 557 AATAGCAATCNCCTATTAGTTAGTGAAGAAACAGATCTCAACAGAGTTTGTGACAG 616
QY 541 ACCGACAGAAAAAGTGGAACTGTGCTGTGCTTGAAGAAAGCGCTGACCAAGCG 600
Db 617 ACCGACAGAAAAAGTGGAACTGTGCTGTGCTTGAAGAAAGCGCTGACCAAGCG 676
QY 601 TTCCCAAGGGGCGAGCTCTTCCGCGCAACCGACTGCTGAGTTTCCGGGTTTCT 660
Db 677 TTCCCAAGGGGCGAGCTCTTCCGCGCAACCGACTGCTGAGTTTCCGGGTTTCT 736
QY 661 AAGACTCTGAGCTGTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 720
Db 737 AAGACTCTGAGCTGTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 796
QY 721 CCCCTGAGCGCTCTCTAAGAGGGGGGCGCGGAGCCCGGCGAGCAAGAAAGAGCG 780
Db 797 CCCCTGAGCGCTCTCTAAGAGGGGGGCGCGGAGCCCGGCGAGCAAGAAAGAGCG 856
QY 781 GAGCGGAGGACGGCGCGGAGAAAGCGCGGAGAAAGGGTCTGTCACACCGGCACTTGG 840
Db 857 GAGCGGAGGACGGCGCGGAGAAAGCGCGGAGAAAGGGTCTGTCACACCGGCACTTGG 916
QY 841 CCTGCTCGCTCCGCGCGGCGCACTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 900
Db 917 CCTGCTCGCTCCGCGCGGCGCACTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 976
QY 901 CCTGCTCGCTCCGCGCGGCGCACTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 960
Db 977 CCTGCTCGCTCCGCGCGGCGCACTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1036
QY 961 TTGCGCCCGCGCGCAAGAGAGGGAGCGCGAGCCCGGCTGCGCGCAAGGGAACTTGG 1020
Db 1037 TTGCGCCCGCGCGCAAGAGAGGGAGCGCGAGCCCGGCTGCGCGCAAGGGAACTTGG 1096
QY 1021 TACCTGGGAGAGATCAACAACATCTCTCGCTGGGGGTGAGAGAAAGAGCAACCGGAG 1080
Db 1097 TACCTGGGAGAGATCAACAACATCTCTCGCTGGGGGTGAGAGAAAGAGCAACCGGAG 1156
QY 1081 ACCGACACCTGTGCTGATTTGGGCAATGAGAGCGGCTGACAGGCTGAGAGTGAAGCGCG 1140
Db 1157 ACCGACACCTGTGCTGATTTGGGCAATGAGAGCGGCTGACAGGCTGAGAGTGAAGCGCG 1216
QY 1141 GGCCCTGCGGAGCGGGTGGCGGAGAGAGGAGCGCGGCTGGGGA 1187

Db 1217 GGCCCTGCGGAGCGGGTGGCGGAGAGAGGAGCGCGGCTGGGGA 1263

RESULT 3
US-10-257-166-172/c
Sequence 172, Application US/10257166
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Method and Nucleic Acids for Analysing the Methylation of
TITLE OF INVENTION: Gene Implicated in Pharmacogenomics
FILE REFERENCE: 5013.1011
CURRENT APPLICATION NUMBER: US/10/257,166
CURRENT FILING DATE: 2002-10-07
PRIOR APPLICATION NUMBER: PCT/EP01/07470
DE 10032529.7
DE 10043826.1
PRIOR FILING DATE: 2001-06-29
2000-06-30
2000-09-01
NUMBER OF SEQ ID NOS: 178
SEQ ID NO 172
LENGTH: 1186
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-257-166-172

Query Match 64.9%; Score 770.6; DB 6; Length 1186;
Best Local Similarity 78.6%; Pred. No. 3.2e-96;
Matches 928; Conservative 0; Mismatches 252; Indels 1; Gaps 1;

QY 2 ATGCGGCACTGACCTCCAGGCTGGGTGAGAGGCGAGCTGCTGCTCAAAAAA 61
Db 1185 ATGCGGCACTGACCTCCAGGCTGGGTGAGAGGCGAGCTGCTGCTCAAAAAA 1126
QY 62 AAAAGACCGCCAGGCTCAAAACAAAAAAGCTTGGAAAAAGCTTGGGCTTTT 121
Db 1125 AAAAGACCGCCAGGCTCAAAACAAAAAAGCTTGGAAAAAGCTTGGGCTTTT 1066
QY 122 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 181
Db 1065 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 1066
QY 182 GGATCTGCTCACTGCAACCTTGGCTCCAGGTTCAAGCAATCTTCTGCTGAGCT 241
Db 1005 GGATCTGCTCACTGCAACCTTGGCTCCAGGTTCAAGCAATCTTCTGCTGAGCT 946
QY 242 CCCAAGTAGCCACACGCCAGCTAATTTTGTATTTTGTATAGAGAGGGGTTTCA 301
Db 945 CCCAAGTAGCCACACGCCAGCTAATTTTGTATTTTGTATAGAGAGGGGTTTCA 886
QY 302 ATGTTGCAAGGCTGTGAGCACTGCTGACCTCAAGTATCAACCGGCTGAGCCGCC 361
Db 885 ATGTTGCAAGGCTGTGAGCACTGCTGACCTCAAGTATCAACCGGCTGAGCCGCC 826
QY 362 AAAGTACTAGATTTACAGGCTGAGGACCGGCTCCAGGCTGAGGCTGAGGCTTGA 421
Db 825 AAAGTACTAGATTTACAGGCTGAGGACCGGCTCCAGGCTGAGGCTTGA 766
QY 422 GTAGAAAAGCTGATTAATACACTTGGCTTGGTGGCTTGGTGAAGAGAAATGGA 481
Db 765 GTAGAAAAGCTGATTAATACACTTGGCTTGGTGGCTTGGTGAAGAGAAATGGA 707
QY 482 AATGCAATCTTATTTGTTGAGAAAGATCTCAACAGAGCTTGTGAGCAAGAG 541
Db 706 AATGCAATCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 647
QY 542 CCGAGGAAAGCTGAGAACTGTGCTGCTGCTTGAAGAGGCGGCTGCAAGAGCGGT 601


```

RESULT 4
US-10-257-166-171
/ Sequence 171, Application US/10257166
/ GENERAL INFORMATION:
/ APPLICANT: OLEK, Alexander
/ APPLICANT: PIEPENBROCK, Christian
/ APPLICANT: BERLIN, Kurt
/ TITLE OF INVENTION: Method and Nucleic Acids for Analysing the Methylation of
/ FILE REFERENCE: 5013.1011
/ CURRENT APPLICATION NUMBER: US/10/257,166
/ CURRENT FILING DATE: 2002-10-07
/ PRIOR APPLICATION NUMBER: PCT/EP01/07470
/ DE 10032529.7
/ DE 10043826.1
/ PRIOR FILING DATE: 2001-06-29
/ 2000-06-30
/ 2000-09-01
/ NUMBER OF SEQ ID NOS: 178
/ SEQ ID NO 171
/ LENGTH: 1186
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
/ US-10-257-166-171

```

59.8%; Score 710; DB 6; Length 1186;
75.1%; Pred. No. 4.5e-88;

| | | | |
|----|------|--|------|
| QY | 1 | GATCGCGCATCTGACATCCAGACCTCGGAGAGAGAGAGAGAGATCTGCTCAAAAAA | 60 |
| Db | 1 | GATCGCGCATTAATGTAATTAAGTTTGGGAGAGAGAGACAGATTTTGTAAAAA | 60 |
| QY | 61 | AAAAAGACCGGACAGGGCTCAAAACAATAAAGCTCGGAAAAAGCCCTGGCGTCTTT | 120 |
| Db | 61 | AAAAAGATCGTTAGGGGTTAAATAAAAATTTCCGAAAAAGTTTGGCGGTTTT | 120 |
| QY | 121 | TTTTTTTTTTTTTTTTTTTTTGGGACAGCTTTCCTCGTCGTCGCCAGCGTGAAGT | 180 |
| Db | 121 | TTTTTTTTTTTTTTTTTTTTTGGGAGTAGTTTGTTCGTATAGGTGAGTAATAGT | 180 |
| QY | 181 | CGAGCTTGCTCACTGCAACTGCTCCAGGTTCAAGCAATTTCTTGCTCAAGCC | 240 |
| Db | 181 | CGAGTTTGGTTATGTGTAATTTGCTTTTAAAGTTAAGTAAATTTTGTGTAGTT | 240 |
| QY | 241 | TCCAAATGACCAACAGCCAGCTAATTTTGTATTTTATAGTAGAGACGGGGGTTTAC | 300 |
| Db | 241 | TTTAAATGATTAATTAAGTTAGTTAAATTTTGTATTTTAAAGTAGAGAGGGGTTTAT | 300 |
| QY | 301 | CATGTTTCCAGGCTGGCTGTCNAACTCCGACCTCGAGGTATCCACCGCTCGGACCC | 360 |
| Db | 301 | TATGTTGTTAGGTGGTTGGTTGAATTTTGAATTTAGGTAGATTTATTCGTTCCGTT | 360 |
| QY | 361 | CAAAATGACTAGATTTAACAAGCGTAGACCAACCGCGTCCAGCGCCCTGGCGGTTTATCA | 420 |
| Db | 361 | TAAAGATTAGGATTAATAGGGGTAGATTAAGCGTTTACGCTTTAGCGGTTTTTAAATTA | 420 |
| QY | 421 | AGTAGAAAAGCTGCATATACCACTTGCTTCNGTTCNTTCAGTAGAAAGAAATATG | 480 |
| Db | 421 | AGTAGAAAAGCTGTTATTAATTTGTTGGTTG - TTTAATGAGAAACGAAGAAATG | 479 |
| QY | 481 | AAATGCAATTCNTATTAAGTTAGTAGAAACAGATCTCAACAGCTTTTGTATAG | 540 |
| Db | 480 | AAATGTAATTTTATTAATGTTGTAAGAAATAGATTTAAATAGTAGTTTGTATAG | 539 |
| QY | 541 | ACCGAGAAAACGTGGAACTGTGCTGCTGCTTAAAGAAAGGCGCGTGCACACAGAG | 600 |
| Db | 540 | ATCGTAGAAAACGTGGAAATGTGTTGTTAGTAGAAAGGCGCGTGCATTAAGAG | 599 |
| QY | 601 | TTTCCAAAGGGCGAGCTCTCCNCGCACAGCACTCGTCCAGAGTCCCGGATTCCT | 660 |
| Db | 600 | TTTTTAAAGGGGTAGTTTATTAAGTTATCGTATTTGTAATTAAGTTTGGGTTTT | 659 |
| QY | 661 | AAGACTCTCAGCGTGAAGCCCGGAGCTCCGTTGTCAGACAACCGTGGCTCTCGGTTTC | 720 |
| Db | 660 | AAGATTTTAAAGTTAGTTGGTTTGGGTTTCGTTTGTGTTAATTCGTGAGTTTGGGTTTT | 719 |
| QY | 721 | CCCTGCGCAGCCTCTATAGAGCGGGGCGCGCAGACCCCGCGAGCAGAAAGAGCG | 780 |
| Db | 720 | TTTTTTGACGACGTTTTTTTAAAGCGGGGCGTCCGGAATTCGTGAGTAGAAGAGCG | 779 |
| QY | 781 | GAGCGCGGACCGCGCGCGGAAAGGCGCGCGAGAGGGGCTCTGACCGCGCATTTGG | 840 |
| Db | 780 | GAGCGCGGACCGTCCGCGGAAAGGCGCGCGGAGGGGTTTGTATCGCTTAATTTGG | 839 |
| QY | 841 | CGTGCCTCGTCCGCGCGCGGCACTTGCTGCTCCGCGCGCGGCACTTGCGCTG | 900 |
| Db | 840 | TTTGTGTTTCGTTTCGTGCGGTTATTTGTTGTTGTTTTCGTTTCGTGTTATTCGTTGG | 899 |
| QY | 901 | CCTCGTCCCGCGCGCGCGCGCAATGCTGTGAGCGGCTCGAGAGCTCGCGCGGCC | 960 |
| Db | 900 | TTTTTGTTTTTCGTTTCGTGCGGTTATGTTGTGTCGTTCGAGTTGTGCGCTGCGTTT | 959 |
| QY | 961 | TTGCCCCCGCGCACAGAGCGGGAAGCGGACCGCGCTCCGCGCACCGGAGAGCTGACG | 1020 |
| Db | 960 | TTGTTTTCGTCGTATAGAGGAGGAGAGTGCAGATCCGTTTCGTGTAAGGAGGTTTGA | 1019 |
| QY | 1021 | TACCGGGGAGATTCACAACATCTCCGTCGCGGCGTCAAGAGAGAGACCGCACCGGCG | 1080 |
| Db | 1020 | TATTTGGGGTAGATTAATAATTTTTCGTGTCGCGGCTTAAAGAGACGATCTGACGGT | 1079 |

| Qy | 1081 | ATGGGAGACCCCTGTGGATTTCCGGATACGACAGGCGCTACACGCAAGGTGACGCGCG | 1140 |
|----|------|--|------|
| Db | 1080 | ATGGGATATTTTGTCCGATTTCCGATATGTATGACCCCGTATAGTTTGAAGGTGACGTCGCG | 1139 |
| Qy | 1141 | GGCCCTGTGCGGAGCGGTGCGCGGAGAGGAGAGGCGCGCTCGGCGA | 1187 |
| Db | 1140 | GGTTTTTGGCGGACGGGTGCGCGGAGAGGAGAGGCGCGCTCGGCGA | 1186 |

```

RESULT 5
US-09-724-676-34523/C
: Sequence 34523, Application US/09724676
: GENERAL INFORMATION:
: APPLICANT: Comugen LTD
: TITLE OF INVENTION: Variants of alternative splicing
: FILE REFERENCE: 129481.4 Comugen
: CURRENT APPLICATION NUMBER: US/09/724,676
: CURRENT FILING DATE: 2000-11-28
: NUMBER OF SEQ. ID NOS: 97222
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 34523
: LENGTH: 1942
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-724-676-34523

```

```
Query Match      25.4%; Score 301; DB 5; Length 1942;
Best Local Similarity 100.0%; Pred. No. 4.6e-33;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

| | | | |
|----|------|---|------|
| Qy | 833 | CAACCTTGGCTGCTCCCGTCCCGGCGGCGACACTTGGACCTGGCTCCGTTCCCGGCGGCGAC | 892 |
| Db | 1329 | CAACTTGGGCTTCCGTTCCGTTCCCGCGGCGCACTTGGACCTGGCTCCGTTCCCGGCGGCGAC | 1870 |
| Qy | 893 | TTGGCTGCTCCGTTCCCGGCGGCGGCGGCGCATCTGTGGCGGCTCGGAGACTGCGCG | 952 |
| Db | 1869 | TTGGCTGCTCCGTTCCCGGCGGCGGCGGCGCATCTGTGTGGCGGCTCGGAGACTGCGCG | 1810 |
| Qy | 953 | GCGGCGCTTTGGCTCCCGGCGGCGCAAGAGAGGGAAGCGAGCGCGTCCGCGCGCAAGGAG | 1012 |
| Db | 1809 | GCGGCGCTTTGGCTCCCGGCGGCGCAAGAGAGGGAAGCGAGCGCGTCCGCGCGCAAGGAG | 1750 |
| Qy | 1013 | AGCTGACGTACTGGGCGAGATCCAAACAATCTCTCGCTGCGGCGGTCCAGGAGAGCGAAC | 1072 |
| Db | 1749 | AGCTGACGTACTGGGCGAGATCCAAACAATCTCTCGCTGCGGCGGTCCAGGAGAGCGAAC | 1690 |
| Qy | 1073 | GCAAGGCGCAACGGGCACTCTGTCCGTAATTGGCATGCAAGCGCGCTTACAGCTTGAAGGTG | 1132 |
| Db | 1689 | GCAAGGCGCAACGGGCACTCTGTCCGTAATTGGCATGCAAGCGCGCTTACAGCTTGAAGGTG | 1630 |
| Qy | 1133 | A 1133 | |
| Db | 1629 | A 1629 | |

```

RESULT 6
US-09-724-676A-34523/C
; Sequence 34523, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34523
; LENGTH: 1942
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-34523

```

Query Match 25.4%; Score 301; DB 5; Length 1942;

Best Local Similarity 100.0%; Pred. No. 4.6e-33;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|------|---|------|
| Qy | 833 | CCACTTGGCCTGCTCCGCTCCGCGCGCCACCTTGGCCTGCTCCGCTCCGCGCGCGCCAC | 892 |
| Db | 1293 | CCACTTGGCCTGCTCCGCTCCGCGCGCCACCTTGGCCTGCTCCGCTCCGCGCGCGCCAC | 1870 |
| Qy | 893 | TTGCGCTGCTCCGCTCCGCGCGCCACCTTGGCCTGCTCCGCTCCGCGCGCGCCAC | 952 |
| Db | 1869 | TTGCGCTGCTCCGCTCCGCGCGCCACCTTGGCCTGCTCCGCTCCGCGCGCGCCAC | 1811 |
| Qy | 953 | GCGGCGCTTGGCTCCGCGCGCCACCTTGGCCTGCTCCGCTCCGCGCGCGCCAC | 1012 |
| Db | 1809 | GCGGCGCTTGGCTCCGCGCGCCACCTTGGCCTGCTCCGCTCCGCGCGCGCCAC | 1756 |
| Qy | 1013 | AGCTGAGTACCTGGGAGATCCAAACATCTCTGCGGAGGTAGGAAGAGACAC | 1072 |
| Db | 1749 | AGCTGAGTACCTGGGAGATCCAAACATCTCTGCGGAGGTAGGAAGAGACAC | 1690 |
| Qy | 1073 | GCAAGGAGCAGGAGCAGCTTGGCTGCTGAGTTCGCGATGCAAGGAGGCTGAGAGGTG | 1132 |
| Db | 1689 | GCAAGGAGCAGGAGCAGCTTGGCTGCTGAGTTCGCGATGCAAGGAGGCTGAGAGGTG | 1630 |
| Qy | 1133 | A | |
| Db | 1629 | A | |

| | | | | |
|---------------------------|---------|--------------------|-------|-------------------|
| Query Match | 25.4%; | Score 301; | DB 5; | length 2042; |
| Best Local Similarity | 100.0%; | Pred. No. 4,6e-33; | | |
| Matches 301; Conservative | 0; | Mismatches | 0; | Indels 0; Gaps 0; |

| | | | |
|----|-------------|---|------|
| Qy | 833 | CCACGTGGACCTGCCTCCGTCGCCGCGGCACTTGAGCTGCTCCGTCGCCGCGCAC | 892 |
| Db | 2029 | CCACTTGGCTGCTCCTCCGTCGCCGCGGCACTTGAGCTGCTCCGTCGCCGCGCAC | 1970 |
| Qy | 893 | TTCCGCTGCTGCCTGCCGCCGCCGCGGCACTGCTGAGCGGACTCGAGAGTGCAC | 952 |
| Db | 1969 | TTCCGCTGCTGCCTCCGTCGCCGCCGCCGCGGCACTGCTGAGCGGACTCGAGAGTGCAC | 1910 |
| Qy | 953 | GCCGCGCCCTTGCCGCCGCCGCCGCGGCAAGAGCGGAGAGCGCCGAGCTCCGCCGCGAGG | 1012 |
| Db | 1909 | GCCGCGCCCTTGCCGCCGCCGCCGCCGCGGCAAGAGCGGAGAGCGCCGAGCTCCGCCGCGAGG | 1850 |
| Qy | 1013 | AGCTGCAGTACTCTGAGGAGCAATCCAAACAATCTCCGCTGCGGCGCTGAGAGGAGACCC | 1072 |
| Db | 1849 | AGCTGCAGTACTCTGAGGAGCAATCCAAACAATCTCCGCTGCGGCGCTGAGAGGAGACCC | 1790 |
| Qy | 1073 | GGAAGGAGCAACCGGCACTCTGTCGGTATTGGGATGCAAGGAGGGCTTACAGCTTGAGAGTG | 1132 |
| Db | 1789 | GGAAGGAGCAACCGGCACTCTGTCGGTATTGGGATGCAAGGAGGGCTTACAGCTTGAGAGTG | 1730 |
| Qy | 1133 A 1133 | | |
| Db | 1729 A 1729 | | |

CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34535
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (41)..(41)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-34535

Query Match 25.4%; Score 301; DB 5; Length 2411;
Best Local Similarity 100.0%; Pred. No. 4.5e-33;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 833 CCACTTGAGCTGCTCCGTCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 892
Db 2398 CCACTTGAGCTGCTCCGTCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 2339
Qy 893 TTGCGCTGCTCCGTCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 952
Db 2338 TTGCGCTGCTCCGTCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 2279
Qy 953 GCCGCGCTTGGCCCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 1012
Db 2278 GCCGCGCTTGGCCCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 2219
Qy 1013 AGCTGAGTACTGAGGAGATCCAAACATCTCCGTCGCGCGGTAGAGAGAGACC 1072
Db 2218 AGCTGAGTACTGAGGAGATCCAAACATCTCCGTCGCGCGGTAGAGAGAGACC 2159
Qy 1073 GCACGGGACCGGCACTTCGCTGATGAGAGCGGTAGAGAGAGAGAG 1132
Db 2158 GCACGGGACCGGCACTTCGCTGATGAGAGCGGTAGAGAGAGAGAG 2099
Qy 1133 A 1133
Db 2098 A 2098

RESULT 12
US-09-724-676A-34535/C
; Sequence 34535, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34535
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (41)..(41)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-34535

Query Match 25.4%; Score 301; DB 5; Length 2411;
Best Local Similarity 100.0%; Pred. No. 4.5e-33;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 833 CCACTTGAGCTGCTCCGTCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 892
Db 2398 CCACTTGAGCTGCTCCGTCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 2339
Qy 893 TTGCGCTGCTCCGTCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 952

Db 2338 TTGCGCTGCTCCGTCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 2279
Qy 953 GCCGCGCTTGGCCCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 1012
Db 2278 GCCGCGCTTGGCCCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 2219
Qy 1013 AGCTGAGTACTGAGGAGATCCAAACATCTCCGTCGCGCGGTAGAGAGAGACC 1072
Db 2218 AGCTGAGTACTGAGGAGATCCAAACATCTCCGTCGCGCGGTAGAGAGAGACC 2159
Qy 1073 GCACGGGACCGGCACTTCGCTGATGAGAGCGGTAGAGAGAGAGAG 1132
Db 2158 GCACGGGACCGGCACTTCGCTGATGAGAGCGGTAGAGAGAGAGAG 2099
Qy 1133 A 1133
Db 2098 A 2098

RESULT 13
US-09-724-676-34589/C
; Sequence 34589, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34589
; LENGTH: 2491
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-34589

Query Match 25.4%; Score 301; DB 5; Length 2491;
Best Local Similarity 100.0%; Pred. No. 4.5e-33;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 833 CCACTTGAGCTGCTCCGTCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 892
Db 2478 CCACTTGAGCTGCTCCGTCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 2419
Qy 893 TTGCGCTGCTCCGTCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 952
Db 2418 TTGCGCTGCTCCGTCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 2359
Qy 953 GCCGCGCTTGGCCCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 1012
Db 2358 GCCGCGCTTGGCCCGCGCGGCACTTGAGCTCCGTCGCGCGCGCAC 2299
Qy 1013 AGCTGAGTACTGAGGAGATCCAAACATCTCCGTCGCGCGGTAGAGAGAGACC 1072
Db 2298 AGCTGAGTACTGAGGAGATCCAAACATCTCCGTCGCGCGGTAGAGAGAGACC 2239
Qy 1073 GCACGGGACCGGCACTTCGCTGATGAGAGCGGTAGAGAGAGAGAG 1132
Db 2238 GCACGGGACCGGCACTTCGCTGATGAGAGCGGTAGAGAGAGAGAG 2179
Qy 1133 A 1133
Db 2178 A 2178

RESULT 14
US-09-724-676A-34589/C
; Sequence 34589, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen

;; CURRENT APPLICATION NUMBER: US/09/724,676A
;; CURRENT FILING DATE: 2000-11-28
;; NUMBER OF SEQ ID NOS: 97222
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 34589
;; LENGTH: 2491
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-724-676A-34589

Query Match 25.4%; Score 301; DB 5; Length 2491;
Best Local Similarity 100.0%; Pred. No. 4.5e-33;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 833 CCACTTGGCTGCTCCGCTCCGCGCCGCGCACTTGGCTGCTCCGCTCCGCGCGCCAC 892
DB 2478 CCACTTGGCTGCTCCGCTCCGCGCCGCGCACTTGGCTGCTCCGCTCCGCGCGCCAC 2419
QY 893 TTGGCTGCTCCGCTCCGCGCCGCGCGCATGCTGTCGCGCTGCGAGCTGCGCGC 952
DB 2418 TTGGCTGCTCCGCTCCGCGCGCGCGCATGCTGTCGCGCTGCGAGCTGCGCGC 2359
QY 953 GCCGCGCTTGCCTCCGCGCGCGCAAGAGCGGAGCGCGCTGCGCGCGCAAGG 1012
DB 2358 GCCGCGCTTGCCTCCGCGCGCGCAAGAGCGGAGCGCGCTGCGCGCGCAAGG 2299
QY 1013 AGCTGCACTGCTGCGCGCGCATGCTGTCGCGCTGCGAGAGGAGCGAGCC 1072
DB 2298 AGCTGCACTGCTGCGCGCGCATGCTGTCGCGCTGCGAGAGGAGCGAGCC 2239
QY 1073 GACGCGGACCGGACCGCTGTCGCTGTCGAGATGACGCGCTGACGCTGAGAGTG 1132
DB 2238 GACGCGGACCGGACCGCTGTCGCTGTCGAGATGACGCGCTGACGCTGAGAGTG 2179
QY 1133 A 1133
DB 2178 A 2178

RESULT 15
US-09-724-676-34548/c
; Sequence 34548, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34548
; LENGTH: 2527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (41)..(41)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-34548

Query Match 25.4%; Score 301; DB 5; Length 2527;
Best Local Similarity 100.0%; Pred. No. 4.5e-33;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 833 CCACTTGGCTGCTCCGCTCCGCGCCGCGCACTTGGCTGCTCCGCTCCGCGCGCCAC 892
DB 2514 CCACTTGGCTGCTCCGCTCCGCGCCGCGCACTTGGCTGCTCCGCTCCGCGCGCCAC 2455
QY 893 TTGGCTGCTCCGCTCCGCGCGCGCATGCTGTCGCGCTGCGAGCTGCGCGC 952
DB 2454 TTGGCTGCTCCGCTCCGCGCGCGCATGCTGTCGCGCTGCGAGCTGCGCGC 2395
QY 953 GCCGCGCTTGCCTCCGCGCGCGCAAGAGCGGAGCGCGCTGCGCGCGCAAGG 1012

DB 2394 GCCGCGCTTGCCTCCGCGCGCGCAAGAGCGGAGCGCGCTGCGCGCGCAAGG 2335
QY 1013 AGCTGCACTGCTGCGCGCGCATGCTGTCGCGCTGCGAGAGGAGCGAGCC 1072
DB 2334 AGCTGCACTGCTGCGCGCGCATGCTGTCGCGCTGCGAGAGGAGCGAGCC 2275
QY 1073 GACGCGGACCGGACCGCTGTCGCTGTCGAGATGACGCGCTGACGCTGAGAGTG 1132
DB 2274 GACGCGGACCGGACCGCTGTCGCTGTCGAGATGACGCGCTGACGCTGAGAGTG 2215
QY 1133 A 1133
DB 2214 A 2214

Search completed: November 29, 2002, 00:20:29
Job time : 115.914 secs

Db 1261 GATCTGACATGCTGAGGTTATGAACAAAGTGAAGATGAATGTATGCTCTTGA 1320
Qy 1321 CAAAAACATGATGTGATTTCAATCCACGACTTATTAAGAAGTGGTGAATTCAC 1380
Db 1321 CAAAAACATGATGTGATTTCAATCCACGACTTATTAAGAAGTGGTGAATTCAC 1380
Qy 1381 AACCTATTTTGGAAATATTTTGAATATTTTGAATTTTCAAGCTATTCCTCAAT 1440
Db 1381 AACCTATTTTGGAAATATTTTGAATATTTTGAATTTTCAAGCTATTCCTCAAT 1440
Qy 1441 CTGAGGAGCTGAGTAAACACCATGATCATGATGTAGAGTGTGTATGAACCTTA 1496
Db 1441 CTGAGGAGCTGAGTAAACACCATGATCATGATGTAGAGTGTGTATGAACCTTA 1496

RESULT 6
PCT-US99-12080-1
Sequence 1, Application PC/TUS9912080
GENERAL INFORMATION:
APPLICANT: Dean, Nicholas M.
TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
TITLE OF INVENTION: HUMAN THYMIDYLATE SYNTHASE EXPRESSION
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Jane Massey Licata
STREET: 66 East Main Street
CITY: Marlton
STATE: NJ
COUNTRY: USA
ZIP: 08053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS 95
SOFTWARE: WORD PERFECT 6.1 FOR WINDOWS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US99/12080
FILING DATE: herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/089,195
FILING DATE: June 2, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: ISPH-0369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 810-1515
TELEFAX: (609) 810-1454
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1536 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Unknown
PUBLICATION INFORMATION:
AUTHORS: Takeishi, K.
AUTHORS: Kaneda, S.
AUTHORS: Aiyasawa, D.
AUTHORS: Shimizu, K.
AUTHORS: Gotoh, O.
AUTHORS: Seno, T.
TITLE: Nucleotide sequence of a functional
TITLE: cDNA for human thymidylate synthase
JOURNAL: Nucleic Acids Res.
VOLUME: 13
ISSUE: 6
PAGES: 2035-2043
DATE: 25-MAR-1985
PCT-US99-12080-1

Query Match 90.8%; Score 1394; DB 1; Length 1536;
Best Local Similarity 99.9%; Pred. No. 0;

Matches 1494; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 GGGGGGGGGGGGGGACCACTTGACCTGCTCCGTCGCGCGGCACTTGACCTGCTCCG 60
Db 1 GGGGGGGGGGGGGGACCACTTGACCTGCTCCGTCGCGCGGCACTTGACCTGCTCCG 60
Qy 61 CCGCGCGGCGCACTTGACCTGCTCCGTCGCGCGGCACTTGACCTGCTCCGTCGCG 120
Db 61 CCGCGCGGCGCACTTGACCTGCTCCGTCGCGCGGCACTTGACCTGCTCCGTCGCG 120
Qy 121 TCGGAGCTGCGCGCGGCTTTCGCGCGGCACTTGACCTGCTCCGTCGCGCGG 180
Db 121 TCGGAGCTGCGCGCGGCTTTCGCGCGGCACTTGACCTGCTCCGTCGCGCGG 180
Qy 181 CCGCGCGGCGGAGCTGCGCGGCACTTGACCTGCGCGGCACTTGACCTGCGCGG 240
Db 181 CCGCGCGGCGGAGCTGCGCGGCACTTGACCTGCGCGGCACTTGACCTGCGCGG 240
Qy 241 AGGAGAGGAGCGGAGCGGCGGCACTTGACCTGCGCGGCACTTGACCTGCGCGG 300
Db 241 AGGAGAGGAGCGGAGCGGCGGCACTTGACCTGCGCGGCACTTGACCTGCGCGG 300
Qy 301 AGCTGAGAGATGAAATTCCTGCTGACCAACCAAGCTGCTGAGAGGCTGTTTG 360
Db 301 AGCTGAGAGATGAAATTCCTGCTGACCAACCAAGCTGCTGAGAGGCTGTTTG 360
Qy 361 GAGGAGCTGCTGCTGCTTTCAGAGGATCCAAATGCTTAAAGCTGCTTTCAGAGG 420
Db 361 GAGGAGCTGCTGCTGCTTTCAGAGGATCCAAATGCTTAAAGCTGCTTTCAGAGG 420
Qy 421 GTTAAATATGAGATGCGCATGATGATCCGAGACTTTTGGACAGCTGAGATTCCTCAC 480
Db 421 GTTAAATATGAGATGCGCATGATGATCCGAGACTTTTGGACAGCTGAGATTCCTCAC 480
Qy 481 AGGAGAGGAGGAGCTTGGGCGGCACTTTCAGAGGAGGAGGAGGAGGAGGAGGAG 540
Db 481 AGGAGAGGAGGAGCTTGGGCGGCACTTTCAGAGGAGGAGGAGGAGGAGGAGGAG 540
Qy 541 TACAGAGATATGATGATGATATTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Db 541 TACAGAGATATGATGATGATATTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Qy 601 GACACCATCAAAACCAACCTGACGACAGAGAGATCATGATGATGATGATGATGAT 660
Db 601 GACACCATCAAAACCAACCTGACGACAGAGAGATCATGATGATGATGATGATGAT 660
Qy 661 GATCTTCTGATGAGGAGCTGCTCCATGCGAGGCTTTCGAGGCTTTCATGATGAG 720
Db 661 GATCTTCTGATGAGGAGCTGCTCCATGCGAGGCTTTCGAGGCTTTCATGATGAG 720
Qy 721 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 AACATGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 781 AACATGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy 841 GGTGACTTTATACACCTTTGGAGATGACATATTTAATCTGAATCAATGAGGACATG 900
Db 841 GGTGACTTTATACACCTTTGGAGATGACATATTTAATCTGAATCAATGAGGACATG 900
Qy 901 AAAATTCAGCTTACGAGAGACCCAGACCTTTCCCAAGGCTCAGAGATTTCTGAAA 960
Db 901 AAAATTCAGCTTACGAGAGACCCAGACCTTTCCCAAGGCTCAGAGATTTCTGAAA 960
Qy 961 GAGAAATTTGATGACTTCAAGCTGAGAGCTTTCAGATGAGAGGATCAATTCGATCA 1020
Db 961 GAGAAATTTGATGACTTCAAGCTGAGAGCTTTCAGATGAGAGGATCAATTCGATCA 1020
Qy 1021 ACTATTAATATGAAATGAGCTGTTTGGGCTGCTTCAAGAGAGCTTGAAGATATTTGCA 1080
Db 1021 ACTATTAATATGAAATGAGCTGTTTGGGCTGCTTCAAGAGAGCTTGAAGATATTTGCA 1080

| | | | |
|----|------|---|------|
| Db | 1021 | ACATTTAAATGGAATTCGCTGTTTAGGGGTCTTCAAGGACCTNAGGAAGATTGTCA | 1080 |
| Qy | 1081 | GTCTTTAGGGGTGGGCTGGATGCCAGGTAAGTCTTTTTGCTTTAAAGANAAG | 1140 |
| Db | 1081 | GTCTTTAGGGGTGGGCTGGATGCCAGGTAAGTCTTTTTGCTTTAAAGANAAG | 1140 |
| Qy | 1141 | AACTAGTCAAAAATCTGTCCTGACCTTCACTTATTAATTTTAAAGATGTGCCAT | 1200 |
| Db | 1141 | AACTAGTCAAAAATCTGTCCTGACCTTCACTTATTAATTTTAAAGATGTGCCAT | 1200 |
| Qy | 1201 | GGCAAAATGAATCTGGCCAGTCTTTTCCATTAATAAAGGCTTGGATTATCATCAGAG | 1260 |
| Db | 1201 | GGCAAAATGAATCTGGCCAGTCTTTTCCATTAATAAAGGCTTGGATTATCATCAGAG | 1260 |
| Qy | 1261 | GGATTCGCAATCTAGGTTATTAACAAGAGGAGAGAAATGATATGCTCTTAG | 1320 |
| Db | 1261 | GGATTCGCAATCTAGGTTATTAACAAGAGGAGAGAAATGATATGCTCTTAG | 1320 |
| Qy | 1321 | CAAAAACATGATGTGCATTTCAATCCACGTACTTAATAAAGGTTGTGAATTCAC | 1380 |
| Db | 1321 | CAAAAACATGATGTGCATTTCAATCCACGTACTTAATAAAGGTTGTGAATTCAC | 1380 |
| Qy | 1381 | AAGCTATTTTGGAAATATTTTGAATATTTTAAAGATTTTCAACAGCTATTCCTCAAT | 1440 |
| Db | 1381 | AAGCTATTTTGGAAATATTTTGAATATTTTAAAGATTTTCAACAGCTATTCCTCAAT | 1440 |
| Qy | 1441 | CTGAGGAGCTGAGTACCACTGATCATTATGTAAGAGTGTGCTATGAACCTTANAAT | 1500 |
| Db | 1441 | CTGAGGAGCTGAGTACCACTGATCATTATGTAAGAGTGTGCTATGAACCTTANAAT | 1500 |
| Qy | 1501 | TGTTTTATGTGCTATPAATAAAGAGTGTCTGC | 1536 |
| Db | 1501 | TGTTTTATGTGCTATPAATAAAGAGTGTCTGC | 1536 |

```

RESULT 5
PCT-US02-18947-556
Sequence 556, Application PC/TUS02/18947
GENERAL INFORMATION:
APPLICANT: Rosetta Impharmatics
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
FILE REFERENCE: 9301-175-228
CURRENT APPLICATION NUMBER: PCT/US02/18947
CURRENT FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 60/380,770
PRIOR FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 2659
SEQ ID NO 556
LENGTH: 1536
TYPE: DNA
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: NM_001071
DATABASE ENTRY DATE: 2001-06-18
PCT-US02-18947-556

```

| | | | | |
|-----------------------|--------------|--------------|---------------|--------------|
| Query Match | 90.8%; | Score 1394; | DB 1; | Length 1536; |
| Best Local Similarity | 99.9%; | Pred. No. 0; | | |
| Matches 1494; | Conservative | 0; | Mismatches 2; | Indels 0; |
| | | | | Gaps 0; |

[illegible]

| | | | |
|----|------|--|------|
| Db | 181 | CCGCCGCAACGGGGAGCTGCAATACCTGGGGAGATCCAAACAATCCTCGCTGGGGGCTC | 240 |
| Qy | 241 | AGGAAGACACCCGCAACGGGCAACCCCTGTGGATTTGGGATCGACGGCCGCTAC | 300 |
| Db | 241 | AGGAAGACACCCGCAACGGGCAACCCCTGTGGATTTGGGATCGACGGCCGCTAC | 300 |
| Qy | 301 | AACCTGAGAGATGAATTCCTCTGCTGACAAACAAACGTGTCTCGAAGGGGTTTGG | 360 |
| Db | 301 | AACCTGAGAGATGAATTCCTCTGCTGACAAACAAACGTGTCTCGAAGGGGTTTGG | 360 |
| Qy | 361 | GAGAGAGTTCGTGTGTTTACAAAGGATCCAAATCTAAAGAGCTGTCTCCAAAGGA | 420 |
| Db | 361 | GAGAGAGTTCGTGTGTTTACAAAGGATCCAAATCTAAAGAGCTGTCTCCAAAGGA | 420 |
| Qy | 421 | GTGAAATCTGGAGTGCATGGAATCCGAGACCTTTTGGACAGCCTGGGATTCACCC | 480 |
| Db | 421 | GTGAAATCTGGAGTGCATGGAATCCGAGACCTTTTGGACAGCCTGGGATTCACCC | 480 |
| Qy | 481 | AGAGAGAGGGGACCTTGGGCCAGATTATGATGCTTCCAGTGGAGGCAATTTGGGGCAGA | 540 |
| Db | 481 | AGAGAGAGGGGACCTTGGGCCAGATTATGATGCTTCCAGTGGAGGCAATTTGGGGCAGA | 540 |
| Qy | 541 | TACAGAGATATGAGATCAGATTTATTCAGAGAAGGGAGTTGACATCAGTCAAGAGATTT | 600 |
| Db | 541 | TACAGAGATATGAGATCAGATTTATTCAGAGAAGGGAGTTGACATCAGTCAAGAGATTT | 600 |
| Qy | 601 | GACACCATCAAAACCAACCTGACACACAGAGATTCATATGTGCGCTTGAATCCAGA | 660 |
| Db | 601 | GACACCATCAAAACCAACCTGACACACAGAGATTCATATGTGCGCTTGAATCCAGA | 660 |
| Qy | 661 | GATCTTCCTCGATAGGGGCGGCTCCATAGCCATAGCCCTCGACAGTTCTATGTGGTAAAC | 720 |
| Db | 661 | GATCTTCCTCGATAGGGGCGGCTCCATAGCCATAGCCCTCGACAGTTCTATGTGGTAAAC | 720 |
| Qy | 721 | AGTGAAGCTTCCTGCTGCACAGCTGTACAGAGATTCGGAGACATGGGCTCTGGTGTGCTTTC | 780 |
| Db | 721 | AGTGAAGCTTCCTGCTGCACAGCTGTACAGAGATTCGGAGACATGGGCTCTGGTGTGCTTTC | 780 |
| Qy | 781 | AACATGCGCAGCTAGGCCCTGCTCAAGTATGATGTGGCAATCAACGGGCTGAAAGCCA | 840 |
| Db | 781 | AACATGCGCAGCTAGGCCCTGCTCAAGTATGATGTGGCAATCAACGGGCTGAAAGCCA | 840 |
| Qy | 841 | GGTGACTTATATCACACTTTGGGAGATGACATATTTACCTGAATACATCGAGCACTGG | 900 |
| Db | 841 | GGTGACTTATATCACACTTTGGGAGATGACATATTTACCTGAATACATCGAGCACTGG | 900 |
| Qy | 901 | AAAAATCAGCTTCAACGAGAACCCCAACCTTTCCCAAAGCTCAGGANTCTTGAAAAATT | 960 |
| Db | 901 | AAAAATCAGCTTCAACGAGAACCCCAACCTTTCCCAAAGCTCAGGANTCTTGAAAAATT | 960 |
| Qy | 961 | GAGAAAAATGAGCTTCAAAAGCGAAGACTTCAGATTGAAGGGATCAATCCGATCCA | 1020 |
| Db | 961 | GAGAAAAATGAGCTTCAAAAGCGAAGACTTCAGATTGAAGGGATCAATCCGATCCA | 1020 |
| Qy | 1021 | ACTATTAATAATGGAATAGCGCTTTTAAAGGCTTCAAGAGCTGTAAGATATTGTCA | 1080 |
| Db | 1021 | ACTATTAATAATGGAATAGCGCTTTTAAAGGCTTCAAGAGCTGTAAGATATTGTCA | 1080 |
| Qy | 1081 | GTCTTTAGGGGTGGGCTGATGTCCGAGGTAAAGTTCTTTTGTCTTAAGAAAAAAGG | 1140 |
| Db | 1081 | GTCTTTAGGGGTGGGCTGATGTCCGAGGTAAAGTTCTTTTGTCTTAAGAAAAAAGG | 1140 |
| Qy | 1141 | AACATGCTAAAAATCTGTCCGTGACCTATCATGATTATTAATTTTAAAGATGTGCCACT | 1200 |
| Db | 1141 | AACATGCTAAAAATCTGTCCGTGACCTATCATGATTATTAATTTTAAAGATGTGCCACT | 1200 |
| Qy | 1201 | GCGAAATGTACTGTGCGCATGTTCTTTTCCATATTAAGAGCTTTGATTAATCATCTAGAG | 1260 |
| Db | 1201 | GCGAAATGTACTGTGCGCATGTTCTTTTCCATATTAAGAGCTTTGATTAATCATCTAGAG | 1260 |
| Qy | 1261 | GTATCTGACATGCTGAGGTTATGAACAAAGTGAGAGAAAGAAATGATATGTGCTTAA | 1320 |

Qy 1081 GCTTTAGGGGTTGGCTGATGCGAAGTAAGTCTTTTCTCTCTTAAGAAAG 1140
 Db 1081 GCTTTAGGGGTTGGCTGATGCGAAGTAAGTCTTTTCTCTTAAGAAAG 1140
 Qy 1141 AACTAGTCAAAAATCTGTCGACCTATCAATTATTAATTTTAAGATGTCAC 1200
 Db 1141 AACTAGTCAAAAATCTGTCGACCTATCAATTATTAATTTTAAGATGTCAC 1200
 Qy 1201 GGCATATGTAATCTGCGAGTCTTTCTCAATAAAGCTTTAGATTAATCACTAG 1260
 Db 1201 GGCATATGTAATCTGCGAGTCTTTCTCAATAAAGCTTTAGATTAATCACTAG 1260
 Qy 1261 GTATCGACAATGCTGAGTTATGAACAAGTGAAGTAAGTAAGTATGCTCTAG 1320
 Db 1261 GTATCGACAATGCTGAGTTATGAACAAGTGAAGTAAGTAAGTATGCTCTAG 1320
 Qy 1321 CAAATATGTAATCTGCTTCAATCCACGACTTATTAAGAGTGTGTAATTCAC 1380
 Db 1321 CAAATATGTAATCTGCTTCAATCCACGACTTATTAAGAGTGTGTAATTCAC 1380
 Qy 1381 AAGCTATTTTGAATATTTTGAATATTTTGAATTTTCAAGCTATTCCTCAAT 1440
 Db 1381 AAGCTATTTTGAATATTTTGAATATTTTGAATTTTCAAGCTATTCCTCAAT 1440
 Qy 1441 CTGAGGAGCTGATGACACCATGATCATGATGATGATGATGATGATGATGAT 1500
 Db 1441 CTGAGGAGCTGATGACACCATGATCATGATGATGATGATGATGATGATGAT 1500
 Qy 1501 TGTATATGTTGCTATTAATAAGAGTCTCTG 1536
 Db 1501 TGTATATGTTGCTATTAATAAGAGTCTCTG 1536

RESULT 4

US-09-963-333-6

Sequence 6, Application US/09963333

GENERAL INFORMATION:

APPLICANT: Stanton, Jr., Vincent P.

TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES

TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT

TITLE OF INVENTION: OF DISEASE

FILE REFERENCE: 11926-015002

CURRENT APPLICATION NUMBER: US/09/963,333

CURRENT FILING DATE: 2001-09-24

PRIOR APPLICATION NUMBER: 09/658,659

PRIOR FILING DATE: 2000-09-08

PRIOR APPLICATION NUMBER: 09/596,033

PRIOR FILING DATE: 2000-06-15

PRIOR APPLICATION NUMBER: 09/357,743

PRIOR FILING DATE: 1999-07-20

PRIOR APPLICATION NUMBER: 09/357,024

PRIOR FILING DATE: 1999-07-19

PRIOR APPLICATION NUMBER: 60/093,484

PRIOR FILING DATE: 1998-07-20

NUMBER OF SEQ ID NOS: 16

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6

LENGTH: 1536

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: 1066

OTHER INFORMATION: n = t or c

NAME/KEY: misc feature

LOCATION: 1136

OTHER INFORMATION: n = a or g

NAME/KEY: misc feature

LOCATION: 1497

OTHER INFORMATION: n = t or a

US-09-963-333-6

Query Match 100.0%; Score 1536; DB 36; Length 1536;

Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGCACTTGGCTGCTCTGCTCCGCGCGGCACTTGGCTGCTCGT 60
 Db 1 GGGGGGGGGGGGCACTTGGCTGCTCTGCTCCGCGCGGCACTTGGCTGCTCGT 60
 Qy 61 CCGCGCGGCACTTGGCTGCTCTGCTCCGCGCGGCACTTGGCTGCTCGT 120
 Db 61 CCGCGCGGCACTTGGCTGCTCTGCTCCGCGCGGCACTTGGCTGCTCGT 120
 Qy 121 TCGAGAGTCCGCGCGGCTTGGCTGCTCTGCTCCGCGCGGCACTTGGCTGCT 180
 Db 121 TCGAGAGTCCGCGCGGCTTGGCTGCTCTGCTCCGCGCGGCACTTGGCTGCT 180
 Qy 181 CCGCGCGGCACTTGGCTGCTCTGCTCCGCGCGGCACTTGGCTGCTCGT 240
 Db 181 CCGCGCGGCACTTGGCTGCTCTGCTCCGCGCGGCACTTGGCTGCTCGT 240
 Qy 241 AAG 300
 Db 241 AAG 300
 Qy 301 AAG 360
 Db 301 AAG 360
 Qy 361 GAG 420
 Db 361 GAG 420
 Qy 421 GTGAAG 480
 Db 421 GTGAAG 480
 Qy 481 AG 540
 Db 481 AG 540
 Qy 541 TACAG 600
 Db 541 TACAG 600
 Qy 601 GACAG 660
 Db 601 GACAG 660
 Qy 661 GATCTCTCTGATGAG 720
 Db 661 GATCTCTCTGATGAG 720
 Qy 721 AGTGAAGTCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 Db 721 AGTGAAGTCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 Qy 781 AACATGCGAG 840
 Db 781 AACATGCGAG 840
 Qy 841 GGTGAGCTTTTATACACTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 Db 841 GGTGAGCTTTTATACACTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 Qy 901 AAAATTCAGCTTCAAG 960
 Db 901 AAAATTCAGCTTCAAG 960
 Qy 961 GAG 1020
 Db 961 GAG 1020
 Qy 1021 ACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
 Db 1021 ACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080

100

| | | | |
|----|------|---|------|
| QY | 1 | GGGGGGGGGGGGGAAACAATTGGCTTCGCTCCGCTCCGCGCGCAACATTGGACCTGGCTCCGG | 60 |
| Db | 1 | GGGGGGGGGGGGGAAACAATTGGCTTCGCTCCGCTCCGCGCGCAACATTGGACCTGGCTCCGG | 60 |
| QY | 61 | CCCCCGCGGCAACTTGGCGCTGCCTCCGCTCCCGCGCGCGCATGCTGTGGCGCGAC | 120 |
| Db | 61 | CCCCCGCGGCAACTTGGCGCTGCCTCCGCTCCCGCGCGCGCATGCTGTGGCGCGAC | 120 |
| QY | 121 | TGGAGATGTCGGCGCGCTTCGCTTCGCCCCCGCGCGCAAGGAGCGGGAGCGCGAGCGCGT | 180 |
| Db | 121 | TGGAGATGTCGGCGCGCTTCGCTTCGCCCCCGCGCGCAAGGAGCGGGAGCGCGAGCGCGT | 180 |
| QY | 181 | CCCGCGCAACGGGGAGCTGCAGTACCTTGGGGGAGATCCAAACAATCTCTCGCTCGCGCTC | 240 |
| Db | 181 | CCCGCGCAACGGGGAGCTGCAGTACCTTGGGGGAGATCCAAACAATCTCTCGCTCGCGCTC | 240 |
| QY | 241 | AGGAGAGACGACCCGACCGGGCACCGGCACTCTGTGGTATTCCGCATTCGAGGCGCTTAC | 300 |
| Db | 241 | AGGAGAGACGACCCGACCGGGCACCGGCACTCTGTGGTATTCCGCATTCGAGGCGCTTAC | 300 |
| QY | 301 | AGCCTGAGAGATGTAATTCCTCTGCTGCAGCAACCAAGTGCTTGAAGGGATGTTTG | 360 |
| Db | 301 | AGCCTGAGAGATGTAATTCCTCTGCTGCAGCAACCAAGTGCTTGAAGGGATGTTTG | 360 |
| QY | 361 | GAGGAGTTCGTGTGGTATTCAAGGGATCCAAATGATCTAAGAGCTGTCTCCAAAGGGA | 420 |
| Db | 361 | GAGGAGTTCGTGTGGTATTCAAGGGATCCAAATGATCTAAGAGCTGTCTCCAAAGGGA | 420 |
| QY | 421 | GTTAAATATGTGGAGATGCACATGAGATCCGAGACTTTTGGACAGCTTGGATTCACAC | 480 |
| Db | 421 | GTTAAATATGTGGAGATGCACATGAGATCCGAGACTTTTGGACAGCTTGGATTCACAC | 480 |
| QY | 481 | AGAGAAAGAGGGGACTTGGGCCAGTTTATGCTTCAATGAGGAGATTTTGGGGCAGAA | 540 |
| Db | 481 | AGAGAAAGAGGGGACTTGGGCCAGTTTATGCTTCAATGAGGAGATTTTGGGGCAGAA | 540 |
| QY | 541 | TACAGAGATATGGAATCAGATTATTCAGAGCAGGGAGTTGACCACTGCAAGAAGATGATT | 600 |
| Db | 541 | TACAGAGATATGGAATCAGATTATTCAGAGCAGGGAGTTGACCACTGCAAGAAGATGATT | 600 |
| QY | 601 | GACACCATCAAAACCACTCTGCAGCAGACAAABAATCATATGTCGCTTGGAAATCAAGA | 660 |
| Db | 601 | GACACCATCAAAACCACTCTGCAGCAGACAAABAATCATATGTCGCTTGGAAATCAAGA | 660 |
| QY | 661 | GATTTTCCTGTATGGCGCTGCTCCATGCGCATGCGCTCTGCGCAAGTTCTATGTGGTAAC | 720 |
| Db | 661 | GATTTTCCTGTATGGCGCTGCTCCATGCGCATGCGCTCTGCGCAAGTTCTATGTGGTAAC | 720 |
| QY | 721 | AGTAGAGTGTCCGTCAGCTGTATACAGAAATCGGGAGCATGGGCTCGAGTGTGCTTTC | 780 |
| Db | 721 | AGTAGAGTGTCCGTCAGCTGTATACAGAAATCGGGAGCATGGGCTCGAGTGTGCTTTC | 780 |
| QY | 781 | AACATGCCACGCTAGCGCTCTCTCAGCTACATGATTGGGCACATCAAGGCGCTTGAAGCA | 840 |
| Db | 781 | AACATGCCACGCTAGCGCTCTCTCAGCTACATGATTGGGCACATCAAGGCGCTTGAAGCA | 840 |
| QY | 841 | GGTATCTTATACACACTTTGGGAGATGCAATATTACCTGAATCAATCGAGCTACTG | 900 |
| Db | 841 | GGTATCTTATACACACTTTGGGAGATGCAATATTACCTGAATCAATCGAGCTACTG | 900 |
| QY | 901 | AAATATCAGCTTCAGGAGAACCCAGACTTTTCCAAAGCTCAGAGATCTTCGAAAAGTT | 960 |
| Db | 901 | AAATATCAGCTTCAGGAGAACCCAGACTTTTCCAAAGCTCAGAGATCTTCGAAAAGTT | 960 |
| QY | 961 | GAGAAATTTATGACTTCAAAAGCTGAAAGACTTTCAGATTGAAGGTTCATTCGCAATCCA | 1020 |
| Db | 961 | GAGAAATTTATGACTTCAAAAGCTGAAAGACTTTCAGATTGAAGGTTCATTCGCAATCCA | 1020 |
| QY | 1021 | ACTATTAATTTGAAATGCGTTTGGGGGCTTTCAAAGGAGCTGTAAGAGATATGTCCA | 1080 |
| Db | 1021 | ACTATTAATTTGAAATGCGTTTGGGGGCTTTCAAAGGAGCTGTAAGAGATATGTCCA | 1080 |

QY 1141 AACCTGCTCAAAATCTGTCGGTACCTATCATGTTATTTTAAAGATGTCACCT 1200
Db 1141 AACCTGCTCAAAATCTGTCGGTACCTATCATGTTATTTTAAAGATGTCACCT 1200
QY 1201 GGCAATGTACTGTGCGAGTCTTTCATATATAAAGGCTTTGAATTACTACTAGG 1260
Db 1201 GGCAATGTACTGTGCGAGTCTTTCATATATAAAGGCTTTGAATTACTACTAGG 1260
QY 1261 GTATCTGACATGCTGAGGTATGAAACAAAGGAGGAAATGAAATGTATGCTCTTAG 1320
Db 1261 GTATCTGACATGCTGAGGTATGAAACAAAGGAGGAAATGAAATGTATGCTCTTAG 1320
QY 1321 CAAAACATGTATGTGATTTTCATCCACGTACTTATAAGAGGTTGTGAAATTTTAC 1380
Db 1321 CAAAACATGTATGTGATTTTCATCCACGTACTTATAAGAGGTTGTGAAATTTTAC 1380
QY 1381 AAGCTATTTTGGATATTTTATTAATAATTTTAAAGATTTTCAAGACTTTCCCTCAAT 1440
Db 1381 AAGCTATTTTGGATATTTTATTAATAATTTTAAAGATTTTCAAGACTTTCCCTCAAT 1440
QY 1441 CTGAGGAGCTGAGTAAACACATCATCATATGATGATGATGATGATGATGATGATG 1500
Db 1441 CTGAGGAGCTGAGTAAACACATCATCATATGATGATGATGATGATGATGATGATG 1500
QY 1501 TGTTTTATATGTGCTATATATAAGAGTGTCTGC 1536
Db 1501 TGTTTTATATGTGCTATATATAAGAGTGTCTGC 1536

RESULT 2

US-09-962-665-6
; Sequence 6, Application US/09962665
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: POLY(POLYGLUTAMATE SYNTHETASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
; FILE REFERENCE: 11926-015004
; CURRENT APPLICATION NUMBER: US/09/962,665
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1536
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1066
; OTHER INFORMATION: n = t or c
; NAME/KEY: misc_feature
; LOCATION: 1136
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc_feature
; LOCATION: 1497
; OTHER INFORMATION: n = t or a
US-09-962-665-6

Query Match 100.0%; Score 1536; DB 36; Length 1536;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGACCACTTGCCCTGCTCCGTCGCCGCGGACCACTTGCCCTGCTCCGCT 60

Db 1 GGGGGGGGGGGGACCACTTGCCCTGCTCCGTCGCCGCGGACCACTTGCCCTGCTCCGCT 60
QY 61 CCGCGCGCGCGCACTTGCGCCGCTCCGTCGCCGCGCGCGCGCGCACTTGCGCGCGCG 120
Db 61 CCGCGCGCGCGCACTTGCGCCGCTCCGTCGCCGCGCGCGCGCGCGCACTTGCGCGCGCG 120
QY 121 TGGAGCTGCG 180
Db 121 TGGAGCTGCG 180
QY 181 CCG 240
Db 181 CCG 240
QY 241 AGAAG 300
Db 241 AGAAG 300
QY 301 AGCTGAG 360
Db 301 AGCTGAG 360
QY 361 GAGGAGGTGCTGCTGCTTATCAAGGATCCACAAATGCTAAAGCTGCTTCCAGAGGA 420
Db 361 GAGGAGGTGCTGCTGCTTATCAAGGATCCACAAATGCTAAAGCTGCTTCCAGAGGA 420
QY 421 GTGAAATCTGGAGATGCAATGATCCGAGACTTTTGTGAGAGAGCTGGAGATCTCCACC 480
Db 421 GTGAAATCTGGAGATGCAATGATCCGAGACTTTTGTGAGAGAGCTGGAGATCTCCACC 480
QY 481 AGAAG 540
Db 481 AGAAG 540
QY 541 TAGAGATATGAGATATGATATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 541 TAGAGATATGAGATATGATATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 601 GAGCAGATCAAAACCAACCTTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 601 GAGCAGATCAAAACCAACCTTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 GATCTTCCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 GATCTTCCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 781 AACATGCGCAGCTGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 781 AACATGCGCAGCTGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 841 GGTACTTTATACACACTTTGGAGATGACATATTTTACTGATATCATGAGAGCACTG 900
Db 841 GGTACTTTATACACACTTTGGAGATGACATATTTTACTGATATCATGAGAGCACTG 900
QY 901 AAAATTCAGCTTACGAGAAACCAAGCTTTCCAAAGCTTCCGAAAGCTTCCGAAAG 960
Db 901 AAAATTCAGCTTACGAGAAACCAAGCTTTCCAAAGCTTCCGAAAGCTTCCGAAAG 960
QY 961 GAGAAATTTGATGATCTTCAAAAGCTGAAAGACTTTCAGATTTGAAAGGATCAATCCG 1020
Db 961 GAGAAATTTGATGATCTTCAAAAGCTGAAAGACTTTCAGATTTGAAAGGATCAATCCG 1020
QY 1021 ACTATTAATAATGAAATGCTGCTTACGCTGCTTCAAAAGAGCTGAAAGATATTTG 1080
Db 1021 ACTATTAATAATGAAATGCTGCTTACGCTGCTTCAAAAGAGCTGAAAGATATTTG 1080
QY 1081 GTCTTAGGGGTTGGGCTGAGTCCGAGGTAAAGTCTTTTGTCTCAAAAGANAAG 1140

22 1382 90.0 1829 29 US-09-726-806-5390 Sequence 5390, Ap
23 1382 90.0 1829 29 US-09-726-807-3895 Sequence 3895, Ap
24 1346 87.6 3298 63 US-60-195-106-200 Sequence 200, App
25 1346 87.6 3298 63 US-60-217-674-28 Sequence 28, App1
26 1346 87.6 3298 71 US-60-278-258-674 Sequence 674, App
27 1280 83.3 1633 61 US-60-172-373-69 Sequence 69, App1
28 1236 80.5 1721 30 US-09-760-475-391 Sequence 391, App
29 1230 80.1 1478 39 US-10-084-817-332 Sequence 332, App
30 942 61.3 942 18 US-09-457-205-29 Sequence 29, App1
31 942 61.3 942 24 US-09-631-275-143 Sequence 143, App
32 938 61.1 1540 22 PCT-US01-14827-3952 Sequence 3952, Ap
33 938 61.1 1540 22 US-09-577-408-4887 Sequence 4887, Ap
34 882 57.4 1395 30 US-09-760-475-1516 Sequence 1516, Ap
35 882 57.4 1395 30 US-09-760-485-46 Sequence 46, App1
36 882 57.4 1395 42 US-10-216-436-46 Sequence 46, App1
37 537 35.0 18597 25 US-09-658-659-8 Sequence 8, App1
38 537 35.0 18597 36 US-09-962-665-8 Sequence 8, App1
39 537 35.0 18597 36 US-09-963-333-8 Sequence 8, App1
40 537 35.0 18597 36 PCT-US01-08631-20764 Sequence 20764, A
41 536 34.9 1539 33 US-09-878-178-2148 Sequence 2148, Ap
42 532 34.6 532 33 US-10-046-935-2148 Sequence 2148, Ap
43 532 34.6 532 38 US-10-146-502-2148 Sequence 2148, Ap
44 532 34.6 532 40 US-09-577-266-11 Sequence 11, App1
45 487 31.7 18596 22

ALIGNMENTS

RESULT 1
US-09-658-659-6

Sequence 6, Application US/09658659

GENERAL INFORMATION:
APPLICANT: Stanton, Jr., Vincent P.
TITLE OF INVENTION: GENE SEQUENCE VARIANCE IN GENES RELATED
TO POLATE METABOLISM HAVING UTILITY IN DETERMINING THE
TITLE OF INVENTION: TREATMENT OF DISEASE
FILE REFERENCE: 11926-015001
CURRENT APPLICATION NUMBER: US/09/658,659
CURRENT FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 09/596,033
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 09/357,743
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 09/357,024
PRIOR FILING DATE: 1999-07-19
PRIOR APPLICATION NUMBER: 60/093,484
PRIOR FILING DATE: 1998-07-20
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 1536
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1066
OTHER INFORMATION: n = t or c
NAME/KEY: misc_feature
LOCATION: 1136
OTHER INFORMATION: n = a or g
NAME/KEY: misc_feature
LOCATION: 1497
OTHER INFORMATION: n = t or a
US-09-658-659-6

Query Match 100.0%; Score 1536; DB 25; Length 1536;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCGGACCACTTGGCTGCTCCGCT 60
Db 1 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCGGACCACTTGGCTGCTCCGCT 60

Qy 61 CCCGCGCGGACCACTTGGCTGCTCCGTCGCGCGGACCACTTGGCTGCTCCGCT 120
Db 61 CCCGCGCGGACCACTTGGCTGCTCCGTCGCGCGGACCACTTGGCTGCTCCGCT 120
Qy 121 TCGAGAGTTCGCGCGCGGCTTGGCTGCTCCGTCGCGGACCACTTGGCTGCTCCGCT 180
Db 121 TCGAGAGTTCGCGCGCGGCTTGGCTGCTCCGTCGCGGACCACTTGGCTGCTCCGCT 180
Qy 181 CCGCGCGGAGGAGTTCGAGTACCTTGGGAGATCCAGATCTCGGCTGGGAGCTC 240
Db 181 CCGCGCGGAGGAGTTCGAGTACCTTGGGAGATCCAGATCTCGGCTGGGAGCTC 240
Qy 241 AGGAGAGGACGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGAC 300
Db 241 AGGAGAGGACGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGAC 300
Qy 301 AGCTGAGAGATGATTCCTCTGCTGACCAACAGTGTCTTGAAGGGTGTGTTG 360
Db 301 AGCTGAGAGATGATTCCTCTGCTGACCAACAGTGTCTTGAAGGGTGTGTTG 360
Qy 361 GAGAGAGTTCGCTGCTTATTCAGAGGATCCAAATGCTTAAAGCTCTTCAAGGA 420
Db 361 GAGAGAGTTCGCTGCTTATTCAGAGGATCCAAATGCTTAAAGCTCTTCAAGGA 420
Qy 421 GTGAATTCGAGATCCGATGATCCGAGATCTTTGGAGAGGATTTGCGAC 480
Db 421 GTGAATTCGAGATCCGATGATCCGAGATCTTTGGAGAGGATTTGCGAC 480
Qy 481 AGAGAGAGGAGGACCTTGGGACCGGACCTTGGGACCTTGGGACCTTGGGAC 540
Db 481 AGAGAGAGGAGGACCTTGGGACCGGACCTTGGGACCTTGGGACCTTGGGAC 540
Qy 541 TACGAGATGATGATCAGATTAATCAGACAGGAGTTCAGCACTGCAAGATGAT 600
Db 541 TACGAGATGATGATCAGATTAATCAGACAGGAGTTCAGCACTGCAAGATGAT 600
Qy 601 GACACATCAAAACCAACCTGACGACGAGAAATCATGTCGCTTGAATCAAGA 660
Db 601 GACACATCAAAACCAACCTGACGACGAGAAATCATGTCGCTTGAATCAAGA 660
Qy 661 GATCTTCCTGATGAGGCTGCTCCATGCAATGCTTGGCACTTGAATGATGAC 720
Db 661 GATCTTCCTGATGAGGCTGCTCCATGCAATGCTTGGCACTTGAATGATGAC 720
Qy 721 AGTGAAGTTCGCTGCTGACGCTGACGAGATCGGAGACATGAGGCTCGGTGCTTTC 780
Db 721 AGTGAAGTTCGCTGCTGACGCTGACGAGATCGGAGACATGAGGCTCGGTGCTTTC 780
Qy 781 AACATCGCAGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 781 AACATCGCAGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Qy 841 GGTGACTTATACCACTTGGGAGTGCATTAATTAATTAATTAATTAATTAATTA 900
Db 841 GGTGACTTATACCACTTGGGAGTGCATTAATTAATTAATTAATTAATTAATTA 900
Qy 901 AAAATTCAGCTTACGAGACCAAGCTTCCCAAGCTGAGATTCGTAAGATTT 960
Db 901 AAAATTCAGCTTACGAGACCAAGCTTCCCAAGCTGAGATTCGTAAGATTT 960
Qy 961 GAGAAATTTGATGACTTCAAGCTGAAAGCTTCAAGTGAAGGATCAATCCGATCA 1020
Db 961 GAGAAATTTGATGACTTCAAGCTGAAAGCTTCAAGTGAAGGATCAATCCGATCA 1020
Qy 1021 ACTATTAATTAATGAAATGCTGTTTGGGCTGTTCAAGAGGCTGAGAGATTTGCA 1080
Db 1021 ACTATTAATTAATGAAATGCTGTTTGGGCTGTTCAAGAGGCTGAGAGATTTGCA 1080
Qy 1081 GTCTTTAGGGGTTGGCTGATGCTGAGGATTAAGTCTTTTGTCTTAAGAAAGAG 1140
Db 1081 GTCTTTAGGGGTTGGCTGATGCTGAGGATTAAGTCTTTTGTCTTAAGAAAGAG 1140

Sun Dec 1 18:15:50 2002

us-09-963-333-6.01igo.rnpn

Page 12

| | | | |
|----|------|--|------|
| Db | 2903 | TCGAGCTCCCGCGCGCGCCCTTTCGCCCGCGCAGCAGAGGCGGAGCGCCAGCGCGCT | 2844 |
| Qy | 181 | CCGCGCCACGCGGAGACTGCAGATACCTTGCGGCGAGATCCAAACATCTCTCGCTCGCGCGCTC | 240 |
| Db | 2843 | CCGCGCCACGCGGAGACTGCAGATACCTTGCGGCGAGATCCAAACATCTCTCGCTCGCGCGCTC | 2784 |
| Qy | 241 | AGGAGAGCGACCGCACGCGGCGACCGCGACCCCTGTGCGATTCGCGCATGCAGGCGCGCTAC | 300 |
| Db | 2783 | AGGAGAGCGACCGCGACCGCGACCCCTGTGCGATTCGCGCATGCAGGCGCGCTAC | 2724 |
| Qy | 301 | AGCCTGAGAGATGATTTCCCTCTGCTGACACAAAGCTGTTCTGAAAGGCTGTTTG | 360 |
| Db | 2723 | AGCCTGAGAGATGATTTCCCTCTGCTGACACAAAGCTGTTCTGAAAGGCTGTTTG | 2664 |
| Qy | 361 | GAGAGCTTCCTGTGTTATCAAGGATCCAAATGCTAAAGAGCTGTTCTCAAGGGA | 420 |
| Db | 2663 | GAGAGCTTCCTGTGTTATCAAGGATCCAAATGCTAAAGAGCTGTTCTCAAGGGA | 2604 |
| Qy | 421 | GTGAAATCTGGGATGCCAATGATCCCGAGACTTTTGGACAAGCTTGCGATTTCTCCACC | 480 |
| Db | 2603 | GTGAAATCTGGGATGCCAATGATCCCGAGACTTTTGGACAAGCTTGCGATTTCTCCACC | 2544 |
| Qy | 481 | AGAGAGAGGCGGACTTGGCGCGAGTTTATGCTTCCAGTGGAGGCACTTTGGGCGCAAA | 540 |
| Db | 2543 | AGAGAGAGGCGGACTTGGCGCGAGTTTATGCTTCCAGTGGAGGCACTTTGGGCGCAAA | 2484 |
| Qy | 541 | TACAGAGATGAGATCCAGTATTCAGAGACAGGAGATGACCAACTGCAGAAAGATGAT | 600 |
| Db | 2483 | TACAGAGATGAGATCCAGTATTCAGAGACAGGAGATGACCAACTGCAGAAAGATGAT | 2424 |
| Qy | 601 | GACACCATCAAAACCAACCTTGACGACAGAAATCATCATGTGCGCTTGGATCCAAAG | 660 |
| Db | 2423 | GACACCATCAAAACCAACCTTGACGACAGAAATCATCATGTGCGCTTGGATCCAAAG | 2364 |
| Qy | 661 | GATCTTCTCTGATGCGCGCTGCTCCATGCGACATGCGCTCTGCGCATGTATGTGCTGAAC | 720 |
| Db | 2363 | GATCTTCTCTGATGCGCGCTGCTCCATGCGACATGCGCTCTGCGCATGTATGTGCTGAAC | 2304 |
| Qy | 721 | AGTGAAGCTGCTGCGCAAGCTGTACACAGATGGGAGACATGGGCTTCGATGCTTTC | 780 |
| Db | 2303 | AGTGAAGCTGCTGCGCAAGCTGTACACAGATGGGAGACATGGGCTTCGATGCTTTC | 2244 |
| Qy | 781 | AACATGCGCAGGTACGCGCGCTGCAGTCAATGATTTGCGCATCACAGCGGCGTGAAGCGA | 840 |
| Db | 2243 | AACATGCGCAGGTACGCGCGCTGCAGTCAATGATTTGCGCATCACAGCGGCGTGAAGCGA | 2184 |
| Qy | 841 | GGTGACTTATATCACATTTGGGAGATGACATATTTACCTGAATTCATCATGAGCGCACTG | 900 |
| Db | 2183 | GGTGACTTATATCACATTTGGGAGATGACATATTTACCTGAATTCATCATGAGCGCACTG | 2124 |
| Qy | 901 | AAATATTCAGCTTCAGCGAGAACCCGACCTTTCCAAAGCTCAGGATTTCTCGAAAAGTT | 960 |
| Db | 2123 | AAATATTCAGCTTCAGCGAGAACCCGACCTTTCCAAAGCTCAGGATTTCTCGAAAAGTT | 2064 |
| Qy | 961 | GAGAGAAATTTGATGACTTCAAAAGCTGAAGACTTTCAGATTTGAAGGATCAATCCGATCCA | 1020 |
| Db | 2063 | GAGAGAAATTTGATGACTTCAAAAGCTGAAGACTTTCAGATTTGAAGGATCAATCCGATCCA | 2004 |
| Qy | 1021 | ACATATTAATATGGAATGCGCTGTTTGGGCGCTTCAAGAGGCTGTAAGAGATTTGTCA | 1080 |
| Db | 2003 | ACATATTAATATGGAATGCGCTGTTTGGGCGCTTCAAGAGGCTGTAAGAGATTTGTCA | 1944 |
| Qy | 1081 | GCTTTTGAAGGATTTGGCGGATGCGAGGATCAAGGATCAAGGATCTTTTCTCT | 1126 |
| Db | 1943 | GCTTTTGAAGGATTTGGCGGATGCGAGGATCAAGGATCAAGGATCTTTTCTCT | 1898 |

Search completed: November 29, 2002, 05:30:50
Job time : 884.958 secs

QY 961 GAGAAATGATGACTTCAAAAGCTGAAAGACTTCAGATTGAAGGGTACAAATCCGATCCA 1020
 DB 1567 GAGAAATGATGACTTCAAAAGCTGAAAGACTTCAGATTGAAGGGTACAAATCCGATCCA 1508
 QY 1021 ACTATTAAATGAAATGGCTGTTTAGGGTCTTTCAAAGAGCTGAGAGATATTGTCA 1080
 DB 1507 ACTATTAAATGAAATGGCTGTTTAGGGTCTTTCAAAGAGCTGAGAGATATTGTCA 1448
 QY 1081 GTCTTTAGGGGTTTGGCTGATGCGGAGTAAAGTCTTTTGTCT 1126
 DB 1447 GTCTTTAGGGGTTTGGCTGATGCGGAGTAAAGTCTTTTGTCT 1402

RESULT 14

US-09-724-676A-34548/c
 ; Sequence 34548, Application US/09724676A
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676A
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 34548
 ; LENGTH: 2527
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (41)..(41)
 ; OTHER INFORMATION: n is a,c,g, or t
 US-09-724-676A-34548

Query Match 66.7%; Score 1024; DB 5; Length 2527;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1124; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGAGACCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGT 60
 DB 2537 GGGGGGGGGGGGAGACCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGT 2468
 QY 61 CCGCGCGCGCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGT 120
 DB 2467 CCGCGCGCGCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGT 2408
 QY 121 TCGAGCTGCGCGCGCTTTCGCTCCGCGCGGCACTTGGCTGCTCCGT 180
 DB 2407 TCGAGCTGCGCGCGCTTTCGCTCCGCGCGGCACTTGGCTGCTCCGT 2348
 QY 181 CCGCGCGCGGAGCTGCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGT 240
 DB 2347 CCGCGCGCGGAGCTGCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGT 2288
 QY 241 AGAAGAGAGACCGGCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGT 300
 DB 2287 AGAAGAGAGACCGGCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGT 2228
 QY 301 AGCTTGAAGATGAATTCCTTCTGCAACAACGCTGCTTCTGAAAGGGTGTGTTG 360
 DB 2227 AGCTTGAAGATGAATTCCTTCTGCAACAACGCTGCTTCTGAAAGGGTGTGTTG 2168
 QY 361 GAGAGATGCTGCTGTTTATCAAGGATCCAAATGCTAAAGAGCTGCTTCCAAAGGGA 420
 DB 2167 GAGAGATGCTGCTGTTTATCAAGGATCCAAATGCTAAAGAGCTGCTTCCAAAGGGA 2108
 QY 421 GTGAAATCTGGAGATGCCAATGATCCCGAGACTTTTGAAGAGCTGCTTCCAAAGGGA 480
 DB 2107 GTGAAATCTGGAGATGCCAATGATCCCGAGACTTTTGAAGAGCTGCTTCCAAAGGGA 2048
 QY 481 AGAAGAGAGGGGACTTGGGCGGCACTTGAAGGCTTCCAGTGAAGGCACTTTTGGGGGAGA 540
 DB 2047 AGAAGAGAGGGGACTTGGGCGGCACTTGAAGGCTTCCAGTGAAGGCACTTTTGGGGGAGA 1988

QY 541 TACAGATATGAAATCAGATTATTTCAGAGCAGGGAGTTGACCAACTGCAAAAGATGATT 600
 DB 1987 TACAGATATGAAATCAGATTATTTCAGAGCAGGGAGTTGACCAACTGCAAAAGATGATT 1928
 QY 601 GACACCATCAAAACCAACCTTGACGACAGAGATCATCATGTGCGTTGAAATCCAGA 660
 DB 1927 GACACCATCAAAACCAACCTTGACGACAGAGATCATCATGTGCGTTGAAATCCAGA 1868
 QY 661 GATCTTCTCTGATGAGGCGCTGCTCCATGCAATGCCCTCTGCGCACTTCAATGAGTGAAC 720
 DB 1867 GATCTTCTCTGATGAGGCGCTGCTCCATGCAATGCCCTCTGCGCACTTCAATGAGTGAAC 1808
 QY 721 AGTACCTGTCTCTGACGCTGTATACAGAGATGGGAGCATGGGCTCTGCTGCTTTC 780
 DB 1807 AGTACCTGTCTCTGACGCTGTATACAGAGATGGGAGCATGGGCTCTGCTGCTTTC 1748
 QY 781 AACATGCGAGCTAGGCGCTCTCAAGTATGATTTGGGCACTACAGGCGCTGAGGCCA 840
 DB 1747 AACATGCGAGCTAGGCGCTCTCAAGTATGATTTGGGCACTACAGGCGCTGAGGCCA 1688
 QY 841 GGTGACTTTAACAACAATTGGGAGATGCAATATTACCTGATCAATCAAGGCCACTG 900
 DB 1687 GGTGACTTTAACAACAATTGGGAGATGCAATATTACCTGATCAATCAAGGCCACTG 1628
 QY 901 AAATTCAGCTTCAAGGAGAACCAACCTTCCCAAGCTGAGATCTTCCAAAAGTT 960
 DB 1627 AAATTCAGCTTCAAGGAGAACCAACCTTCCCAAGCTGAGATCTTCCAAAAGTT 1568
 QY 961 GAGAAATTTGATGACTTCAAGCTGAAAGCTTTCAGATTGAAGGATTCATCCGATCCA 1020
 DB 1567 GAGAAATTTGATGACTTCAAGCTGAAAGCTTTCAGATTGAAGGATTCATCCGATCCA 1508
 QY 1021 ACTATTAAATGAAATGGCTGTTTAGGGTCTTTCAAAGAGCTGAGAGATATTGTCA 1080
 DB 1507 ACTATTAAATGAAATGGCTGTTTAGGGTCTTTCAAAGAGCTGAGAGATATTGTCA 1448
 QY 1081 GTCTTTAGGGGTTTGGCTGATGCGGAGTAAAGTCTTTTGTCT 1126
 DB 1447 GTCTTTAGGGGTTTGGCTGATGCGGAGTAAAGTCTTTTGTCT 1402

RESULT 15

US-09-724-676-34565/c
 ; Sequence 34565, Appl: cation US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 34565
 ; LENGTH: 3023
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-724-676-34565

Query Match 66.7%; Score 1024; DB 5; Length 3023;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1124; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGAGACCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGT 60
 DB 3023 GGGGGGGGGGGGAGACCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGT 2964
 QY 61 CCGCGCGCGCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGT 120
 DB 2963 CCGCGCGCGCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGT 2904
 QY 121 TCGAGCTGCGCGCGCTTTCGCTCCGCGCGGCACTTGGCTGCTCCGT 180

[illegible]

```

; NAME/KEY: misc_feature
; LOCATION: (41).(41)
; OTHER INFORMATION: n ls a,c,g, or t
US-09-724-676-34548

```

| | | | | |
|----------------------------|--------|---------------|-----------|--------------|
| Query Match | 66.7%; | Score 1024; | DB 5; | Length 2527; |
| Best Local Similarity | 99.8%; | Pred. No. 0; | | |
| Matches 1124; Conservative | 0; | Mismatches 2; | Indels 0; | Gaps 0; |

| | | | |
|----|------|--|------|
| QY | 1 | GGGGGGGGGGGGGGGACATTGAGCTGGCTCCGCGCCGGCCGACATTGAGCTGCTCCGT | 60 |
| Db | 2527 | GGGGGGGGGGGGGACCACTTGGCTGGCTCCGCGCCGGCCGACATTGAGCTGCTCCGT | 2468 |
| QY | 61 | CCGGCGGCGCACTTCGCTGCTCCGTCCCGCGCGCGCGCAATGCTGTGGCGGCT | 120 |
| Db | 2467 | CCGGCGGCGCACTTCGCTGCTCCGTCCCGCGCGCGCGCAATGCTGTGGCGGCG | 2408 |
| QY | 121 | TGGAGCTGCGCGCGCGGCTTCCCTCCGCGCAAGAGCGGGAAGCGCAAGCGGT | 180 |
| Db | 2407 | TGGAGCTGCGCGCGCGGCTTCCCTCCGCGCGCAAGAGCGGGAAGCGCAAGCGGT | 2348 |
| QY | 181 | CCCGCGCACGGGAGCTGCACTGTGGGGGAGATCCAAACAATCCCTCCGCTGGCGGCT | 240 |
| Db | 2347 | CCCGCGCACGGGAGCTGCACTGTGGGGGAGATCCAAACAATCCCTCCGCTGGCGGCT | 2288 |
| QY | 241 | AGGAAAGAACGACCGACCGGGACCGGCAACCGTGGATTCCGACATGCAAGCGGCTAC | 300 |
| Db | 2287 | AGGAAAGAACGACCGACCGGGACCGGCAACCGTGGATTCCGACATGCAAGCGGCTAC | 2228 |
| QY | 301 | AGCCTAGAGATGAATTCCTCTGCTGACAAACAACGTGTGTCTGGAAGGATGTTTG | 360 |
| Db | 2227 | AGCCTAGAGATGAATTCCTCTGCTGACAAACAACGTGTGTCTGGAAGGATGTTTG | 2168 |
| QY | 361 | GAGAGCTGCTGTGGTTTATCAAGGATCCAAATGCTAAAGCTGTCTTCCAAAGGA | 420 |
| Db | 2167 | GAGAGCTGCTGTGGTTTATCAAGGATCCAAATGCTAAAGCTGTCTTCCAAAGGA | 2108 |
| QY | 421 | GTGAAATCTGGGATGCCATGATGCCAAGCTTTTGGACAGCTTGGGATTTCTCCAC | 480 |
| Db | 2107 | GTGAAATCTGGGATGCCAATGATGCCAAGCTTTTGGACAGCTTGGGATTTCTCCAC | 2048 |
| QY | 481 | AGAGAGAAAGGGGACTTGGGCCAGTTTATGCTTCCAGTGGAGGCAATTTGGGGGAGAA | 540 |
| Db | 2047 | AGAGAGAAAGGGGACTTGGGCCAGTTTATGCTTCCAGTGGAGGCAATTTGGGGGAGAA | 1988 |
| QY | 541 | TACGAGATTTGGATTCAGATTATTCAAGACGGGAGTTGACCAATCGCAAAAGTGAAT | 600 |
| Db | 1987 | TACGAGATTTGGATTCAGATTATTCAAGACGGGAGTTGACCAATCGCAAAAGTGAAT | 1928 |
| QY | 601 | GAGCAACATCAAAACCAACCGTACGACAGAAAGATCAATGTCGCTTGGATTCACAGA | 660 |
| Db | 1927 | GAGCAACATCAAAACCAACCGTACGACAGAAAGATCAATGTCGCTTGGATTCACAGA | 1868 |
| QY | 661 | GATTTCTCTGATGGCGCTGCTCCATGCAATGCAATCCCTCTGCGAGTTCTATGTGTGAAC | 720 |
| Db | 1867 | GATTTCTCTGATGGCGCTGCTCCATGCAATGCAATCCCTCTGCGAGTTCTATGTGTGAAC | 1808 |
| QY | 721 | AGTAGAGTGTCTTGCAGGTATACAGAGATGGGAGACATGGGCTCTGGTGTGCTTTTC | 780 |
| Db | 1807 | AGTAGAGTGTCTTGCAGGTATACAGAGATGGGAGACATGGGCTCTGGTGTGCTTTTC | 1748 |
| QY | 781 | AACATCGCAGCTAGGCGCTGTCACTATATATTTGGGCACTACAGGGCTGTGAAGCCA | 840 |
| Db | 1747 | AACATCGCAGCTAGGCGCTGTCACTATATATTTGGGCACTACAGGGCTGTGAAGCCA | 1688 |
| QY | 841 | GGTGACTTTATACACTTTGGGAGATGCAATTTTACCTGAATCAATAGAGCCACTG | 900 |
| Db | 1687 | GGTGACTTTATACACTTTGGGAGATGCAATTTTACCTGAATCAATAGAGCCACTG | 1628 |
| QY | 901 | AAATTTGAGCTTAGGGGAACCGAGCTTTTCCCAAACTCAGGATCTTTCGAAAGTT | 960 |
| Db | 1627 | AAATTTGAGCTTAGGGGAACCGAGCTTTTCCCAAACTCAGGATCTTTCGAAAGTT | 1568 |

Db 1391 ACTATTAAATGAAATGGCTGCTTTAGGGTCTTTCAAGAGAGCTCGAAGATATTGTCA 1332
Qy 1081 GTCTTTAGGGGTTTGGGCTGAGATGCCGAGTAAAGTTCTTTTGTCT 1126
Db 1331 GTCTTTAGGGGTTTGGGCTGAGATGCCGAGTAAAGTTCTTTTGTCT 1286

RESULT 11
US-09-724-676-34589/c
; Sequence 34589, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34589
; LENGTH: 2491
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-34589

Query Match 66.7%; Score 1024; DB 5; Length 2491;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1124; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGAGCACTTGGCTGCTTCCGCGCGGCACTTGGCTGCTTCCGT 60
Db 2491 GGGGGGGGGGGGAGCACTTGGCTGCTTCCGCGCGGCACTTGGCTGCTTCCGT 2432
Qy 61 CCGCGCGGCGCACTTGGCTGCTTCCGCGCGGCACTTGGCTGCTTCCGT 120
Db 2431 CCGCGCGGCGCACTTGGCTGCTTCCGCGCGGCACTTGGCTGCTTCCGT 2372
Qy 121 TCGAGAGCTGCGCGCGGCGGCTTGGCGCGGCGGCAAGAGCGGCGGCGT 180
Db 2371 TCGAGAGCTGCGCGCGGCGGCTTGGCGCGGCGGCAAGAGCGGCGGCGT 2312
Qy 181 CCGCGCGGCGGAGCTGAGTACTTGGGAGAGATCCAAATCATCTCGGCTGCGGCGT 240
Db 2311 CCGCGCGGCGGAGCTGAGTACTTGGGAGAGATCCAAATCATCTCGGCTGCGGCGT 2252
Qy 241 AGAAGAGAGCAAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGTAC 300
Db 2251 AGAAGAGAGCAAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGTAC 2192
Qy 301 AGCTTGAAGATGAATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Db 2191 AGCTTGAAGATGAATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2132
Qy 361 GAGGAGTGTGCTGCTTATATCAAGGAGATCCAAATCATCTCTCTCTCTCTCTCT 420
Db 2131 GAGGAGTGTGCTGCTTATATCAAGGAGATCCAAATCATCTCTCTCTCTCTCTCT 2072
Qy 421 GTGAAATCTGGGATGCGCAATGAGATCCGAGACTTTTGGAGAGCTTCTCCACC 480
Db 2071 GTGAAATCTGGGATGCGCAATGAGATCCGAGACTTTTGGAGAGCTTCTCCACC 2012
Qy 481 AGAAGAGAGGAGCTTGGGCGGCAATGAGATCCGAGACTTTTGGAGAGCTTCTCCACC 540
Db 2011 AGAAGAGAGGAGCTTGGGCGGCAATGAGATCCGAGACTTTTGGAGAGCTTCTCCACC 1952
Qy 541 TACAGAGATGAGATCAATTTATCAGAGAGGAGATCCAAATCATCTCTCTCTCTCTCT 600
Db 1951 TACAGAGATGAGATCAATTTATCAGAGAGGAGATCCAAATCATCTCTCTCTCTCTCT 1892
Qy 601 GAGCAGATCAAAACCAACCTTGAAGAGAGAGATCAATGAGAGCTTGGAGATCCAAAGA 660
Db 1891 GAGCAGATCAAAACCAACCTTGAAGAGAGAGATCAATGAGAGCTTGGAGATCCAAAGA 1832
Qy 661 GATCTCTCTGATGGGCGGCTCGATGCGATGCCCTCTGCCAGTTCTTATGTGTGAAC 720

Db 1831 GATCTCTCTGATGGGCGGCTCGATGCGATGCCCTCTGCCAGTTCTTATGTGTGAAC 1772
Qy 721 AGTAGAGCTGCTGCGGAGCTGATACAGAGATCGGAGAGATGAGGCTCGGAGCTTCT 780
Db 1771 AGTAGAGCTGCTGCGGAGCTGATACAGAGATCGGAGAGATGAGGCTCGGAGCTTCT 1712
Qy 781 AACATGCGAGCTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 1711 AACATGCGAGCTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1652
Qy 841 GGTGACTTTATACACTTTTGGAGATGAGATTTTACCTGATATCATCGAGCTGCTG 900
Db 1651 GGTGACTTTATACACTTTTGGAGATGAGATTTTACCTGATATCATCGAGCTGCTG 1592
Qy 901 AAAATTCAGCTTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 1591 AAAATTCAGCTTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1532
Qy 961 GAGAAATTTGATGACTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 1531 GAGAAATTTGATGACTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1472
Qy 1021 ACTATTAAATGAAATGGCTGCTTTAGGGTCTTTCAAGAGAGCTTGAAGATATTGTCA 1080
Db 1471 ACTATTAAATGAAATGGCTGCTTTAGGGTCTTTCAAGAGAGCTTGAAGATATTGTCA 1412
Qy 1081 GTCTTTAGGGGTTTGGGCTGAGATGCCGAGGTAAAGTTCTTTTGTCT 1126
Db 1411 GTCTTTAGGGGTTTGGGCTGAGATGCCGAGGTAAAGTTCTTTTGTCT 1366

RESULT 12
US-09-724-676A-34589/c
; Sequence 34589, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34589
; LENGTH: 2491
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-34589

Query Match 66.7%; Score 1024; DB 5; Length 2491;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1124; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGAGCACTTGGCTGCTTCCGCGCGGCACTTGGCTGCTTCCGT 60
Db 2491 GGGGGGGGGGGGAGCACTTGGCTGCTTCCGCGCGGCACTTGGCTGCTTCCGT 2432
Qy 61 CCGCGCGGCGCACTTGGCTGCTTCCGCGCGGCACTTGGCTGCTTCCGT 120
Db 2431 CCGCGCGGCGCACTTGGCTGCTTCCGCGCGGCACTTGGCTGCTTCCGT 2372
Qy 121 TCGAGAGCTGCGCGCGGCGGCTTGGCGCGGCGGCAAGAGCGGCGGCGGCGT 180
Db 2371 TCGAGAGCTGCGCGCGGCGGCTTGGCGCGGCGGCAAGAGCGGCGGCGGCGT 2312
Qy 181 CCGCGCGGCGGAGCTGAGTACTTGGGAGAGATCCAAATCATCTCTCTCTCTCTCTCT 240
Db 2311 CCGCGCGGCGGAGCTGAGTACTTGGGAGAGATCCAAATCATCTCTCTCTCTCTCTCT 2252
Qy 241 AGAAGAGAGCAAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGTAC 300
Db 2251 AGAAGAGAGCAAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGTAC 2192

Db 2051 GAGAGTTGCTGCTGTTTATCAAGGATCCAAATGCTTAAGAGCTGTTCCAGGGA 1992
QY 421 GTGAAATCTGGATGCGCATGATCCGAGACCTTTTGGACAGCTGGATTTCCACC 480
Db 1991 GTGAAATCTGGATGCGCATGATCCGAGACCTTTTGGACAGCTGGATTTCCACC 1932
QY 481 AGAAGAAGGGGACTTGGGCGGAGTTATGCTTCCAGTGGAGGATTTTGGGCGGAA 540
Db 1931 AGAAGAAGGGGACTTGGGCGGAGTTATGCTTCCAGTGGAGGATTTTGGGCGGAA 1972
QY 541 TACAGATATGATTCAGATTTATTCAGGACAGGGAATTTGACCACTGCAAAAGTAT 600
Db 1871 TACAGATATGATTCAGATTTATTCAGGACAGGGAATTTGACCACTGCAAAAGTAT 1812
QY 601 GACCCATCAAAACCAACCCGTCGACAGAGATCAATCATGTCGCTTGGATTCAGA 660
Db 1811 GACCCATCAAAACCAACCCGTCGACAGAGATCAATCATGTCGCTTGGATTCAGA 1752
QY 661 GATCTTCTGATGGGCTGCTGCTCCATGCGATGCTTCCAGATTTTATGTTGAT 720
Db 1751 GATCTTCTGATGGGCTGCTGCTCCATGCGATGCTTCCAGATTTTATGTTGAT 1692
QY 721 AGTAGCTGTCTGCGACGCTGTAACAGAGATCGGAGACATGGGCTGCTGCTTTC 780
Db 1691 AGTAGCTGTCTGCGACGCTGTAACAGAGATCGGAGACATGGGCTGCTGCTTTC 1632
QY 781 AACATCGCAGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 1631 AACATCGCAGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1572
QY 841 GGTGACTTTTATGACACTTTGGGAGATGACATTTTACTGATTCATGACGCACTG 900
Db 1571 GGTGACTTTTATGACACTTTGGGAGATGACATTTTACTGATTCATGACGCACTG 1512
QY 901 AAAATTCAGCTTACGAGAACCCAGACCTTTCCAAAGCTTCAGATTTTGGAAAGTT 960
Db 1511 AAAATTCAGCTTACGAGAACCCAGACCTTTCCAAAGCTTCAGATTTTGGAAAGTT 1452
QY 961 GAGAAATTTATGATCTTCAAAAGCTGTAAGCTTCAATGATGAGGATTCGATCCA 1020
Db 1451 GAGAAATTTATGATCTTCAAAAGCTGTAAGCTTCAATGATGAGGATTCGATCCA 1392
QY 1021 ACTATTTAAATGGAATGCTGTTTGGGCTTTCAAAGGACTTGAAGGATTTGCA 1080
Db 1391 ACTATTTAAATGGAATGCTGTTTGGGCTTTCAAAGGACTTGAAGGATTTGCA 1332
QY 1081 GTCTTAGGGGTTGGGCTGAGATCGAGGTAAAGTTCTTTTGGCT 1126
Db 1331 GTCTTAGGGGTTGGGCTGAGATCGAGGTAAAGTTCTTTTGGCT 1286

RESULT 10

US-09-724-676A-34535/c
; Sequence 34535, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34535
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (41)..(41)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-34535

Query Match 66.7%; Score 1024; DB 5; Length 2411;

Best Local Similarity 99.8%; Pred. No. 0;
Matches 1124; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGGGGGGGGGGGACACATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Db 2411 GGGGGGGGGGGGACACATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2352
QY 61 CCGGCGGCGACCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db 2351 CCGGCGGCGACCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2292
QY 121 TCGGAGCTGCGCGCGCGCGCTTGGCCCGCGCGACAGAGCGGAGACCGGACCGCT 180
Db 2291 TCGGAGCTGCGCGCGCGCGCTTGGCCCGCGCGACAGAGCGGAGACCGGACCGCT 2232
QY 181 CCGCGGACCGGAGCTGCAATGATCTGGGAGATTCACACATCTCGCTGGGAGGCT 240
Db 2231 CCGCGGACCGGAGCTGCAATGATCTGGGAGATTCACACATCTCGCTGGGAGGCT 2172
QY 241 AGAAGACACCGGACCGGAGACCGGACCGGACCGGACCGGACCGGACCGGACCGCTAC 300
Db 2171 AGAAGACACCGGACCGGAGACCGGACCGGACCGGACCGGACCGGACCGGACCGCTAC 2112
QY 301 AGCCGAGAGATGATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 2111 AGCCGAGAGATGATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2052
QY 361 GAGAGTTGCTGCTGCTTATCAAGGATCCAAATGCTTAAGAGCTGCTTCCAGGGA 420
Db 2051 GAGAGTTGCTGCTGCTTATCAAGGATCCAAATGCTTAAGAGCTGCTTCCAGGGA 1992
QY 421 GTGAAATCTGGATGCGCATGATTCGAGACTTTTGGACAGCTGGGATTTCCACC 480
Db 1991 GTGAAATCTGGATGCGCATGATTCGAGACTTTTGGACAGCTGGGATTTCCACC 1932
QY 481 AGAAGAAGGGGACTTGGGCGGAGTTATGCTTCCAGTGGAGGATTTTGGGCGGAA 540
Db 1931 AGAAGAAGGGGACTTGGGCGGAGTTATGCTTCCAGTGGAGGATTTTGGGCGGAA 1872
QY 541 TACAGATATGATTCAGATTTATTCAGGACAGGGAATTTGACCACTGCAAAAGTAT 600
Db 1871 TACAGATATGATTCAGATTTATTCAGGACAGGGAATTTGACCACTGCAAAAGTAT 1812
QY 601 GACCCATCAAAACCAACCCGTCGACAGAGATCAATCATGTCGCTTGGATTCAGA 660
Db 1811 GACCCATCAAAACCAACCCGTCGACAGAGATCAATCATGTCGCTTGGATTCAGA 1752
QY 661 GATCTTCTGATGGGCTGCTGCTCCATGCGATGCTTCCAGATTTTATGTTGAT 720
Db 1751 GATCTTCTGATGGGCTGCTGCTCCATGCGATGCTTCCAGATTTTATGTTGAT 1692
QY 721 AGTAGCTGTCTGCGACGCTGTAACAGAGATCGGAGACATGGGCTGCTGCTTTC 780
Db 1691 AGTAGCTGTCTGCGACGCTGTAACAGAGATCGGAGACATGGGCTGCTGCTTTC 1632
QY 781 AACATCGCAGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 1631 AACATCGCAGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1572
QY 841 GGTGACTTTTATGACACTTTGGGAGATGACATTTTACTGATTCATGACGCACTG 900
Db 1571 GGTGACTTTTATGACACTTTGGGAGATGACATTTTACTGATTCATGACGCACTG 1512
QY 901 AAAATTCAGCTTACGAGAACCCAGACCTTTCCAAAGCTTCAGATTTTGGAAAGTT 960
Db 1511 AAAATTCAGCTTACGAGAACCCAGACCTTTCCAAAGCTTCAGATTTTGGAAAGTT 1452
QY 961 GAGAAATTTATGATCTTCAAAAGCTGTAAGCTTCAATGATGAGGATTCGATCCA 1020
Db 1451 GAGAAATTTATGATCTTCAAAAGCTGTAAGCTTCAATGATGAGGATTCGATCCA 1392
QY 1021 ACTATTTAAATGGAATGCTGTTTGGGCTTTCAAAGGACTTGAAGGATTTGCA 1080